



**You have downloaded a document from  
RE-BUS  
repository of the University of Silesia in Katowice**

**Title:** Znajdywanie odpowiedniości punktów charakterystycznych na obrazach stereowizyjnych

**Author:** Paweł Popielski

**Citation style:** Popielski Paweł. (2017). Znajdywanie odpowiedniości punktów charakterystycznych na obrazach stereowizyjnych. Praca doktorska. Katowice: Uniwersytet Śląski

© Korzystanie z tego materiału jest możliwe zgodnie z właściwymi przepisami o dozwolonym użytku lub o innych wyjątkach przewidzianych w przepisach prawa, a korzystanie w szerszym zakresie wymaga uzyskania zgody uprawnionego.



UNIwersYTET ŚLĄSKI  
W KATOWICACH



Biblioteka  
Uniwersytetu Śląskiego



Ministerstwo Nauki  
i Szkolnictwa Wyższego

Uniwersytet Śląski  
Wydział Informatyki i Nauki o Materiałach  
Instytut Informatyki

mgr inż. Paweł Popielski

ROZPRAWA DOKTORSKA

ZNAJDYWANIE ODPOWIEDNIOŚCI PUNKTÓW  
CHARAKTERYSTYCZNYCH NA OBRAZACH  
STEREOWIZYJNYCH

Promotor: dr hab. inż. Robert Koprowski  
Promotor wspomagający: dr n. m. Sławomir Wilczyński

Sosnowiec 2017

*Składam serdeczne podziękowania moim promotorom*

**PANU DR. HAB. INŻ. ROBERTOWI KOPROWSKIEMU**

*oraz*

**PANU DR. N. M. SŁAWOMIROWI WILCZYŃSKIEMU**

*za poświęcony czas i pomoc oraz za zaufanie i cierpliwość,  
a także za mobilizację do pracy*

*dziękuję również*

**PANU PROF. DR. HAB. INŻ. ZYGMUNTOWI WRÓBLOWI**

*za cenne wskazówki i mobilizację do pracy*

**PANU DR. ZDZISŁAWOWI KRÓLOWI**

*za udostępnienie klastra obliczeniowego*

**KOLEGOM Z**

*Zakładu Komputerowych Systemów Biomedycznych*

*za stworzenie wspaniałej i niezapomnianej atmosfery pracy*

*szczególnie dziękuję*

**MOIM BLISKIM**

*za cierpliwość, wsparcie duchowe i wiarę we mnie*

## SPIS TREŚCI

<b>1. Cel i zakres pracy .....</b>	<b>6</b>
<b>2. Tezy.....</b>	<b>8</b>
<b>3. Podstawowe definicje i oznaczenia stosowane w pracy .....</b>	<b>10</b>
3.1. Definicje .....	10
3.2. Oznaczenia.....	13
<b>4. Wprowadzenie.....</b>	<b>16</b>
4.1. Rekonstrukcja 3D w medycynie .....	18
4.2. Stereowizja, uzasadnienie wyboru metody.....	19
4.3. Rejestracja.....	21
4.4. Usunięcie dystorsji.....	22
4.5. Elementy orientacji wewnętrznej .....	23
4.6. Elementy orientacji zewnętrznej.....	25
4.7. Rektyfikacja obrazów .....	27
4.8. Poszukiwanie odpowiedniości na obrazach.....	29
4.9. Przestrzeń barwna .....	33
4.9.1. Widzenie barw .....	34
4.9.2. Percepcja barw .....	37
4.9.3. Model barw niezależny od urządzenia wyświetlającego.....	38
4.10. Punkty charakterystyczne .....	40
4.11. Wykrywanie krawędzi .....	43
4.12. Transformata Hough'a.....	46
4.13. Znormalizowana korelacja krzyżowa .....	49
4.14. Metoda wariacyjna.....	51
4.15. Metoda bazująca na informacji wzajemnej. ....	52
4.16. Optymalizacja procesu poszukiwania odpowiedniości .....	53
4.17. Triangulacja chmury punktów .....	55
<b>5. Materiał badawczy .....</b>	<b>57</b>
5.1. Modele stereometryczne .....	58
5.2. Obrazy naczyń podskórnych.....	59
5.3. Rzeczywiste obrazy z Middlebury.....	61
5.4. Macierz dysparycji wzorcowych .....	63
<b>6. Opis metody znajdowania odpowiedniości punktów .....</b>	<b>66</b>

6.1.	Akwizycja obrazów testowych .....	66
6.1.1.	Rejestracja brył .....	66
6.1.2.	Rejestracja naczyń podskórnych .....	67
6.2.	Wstępne przetwarzanie obrazów .....	68
6.2.1.	Segmentacja metodą BGOM .....	69
6.2.2.	Segmentacja metodą SSSB .....	71
6.2.3.	Segmentacja metodą BLG .....	76
6.3.	Pozyskanie cech charakterystycznych ze stereopar .....	77
6.4.	Wybór znaczących cech .....	78
6.5.	Metoda poszukiwania stereokorespondencji MED .....	80
6.5.1.	Analiza złożoności obliczeniowej algorytmu .....	83
6.6.	Metoda poszukiwania stereokorespondencji MED-NDD .....	84
6.6.1.	Analiza złożoności obliczeniowej algorytmu .....	86
6.7.	Metoda poszukiwania stereokorespondencji MED-RGB .....	87
6.7.1.	Analiza złożoności obliczeniowej algorytmu .....	90
6.8.	Przetwarzanie końcowe .....	91
<b>7.</b>	<b>Testowanie opracowanego systemu .....</b>	<b>94</b>
7.1.	Ilościowa ocena wyników poszukiwania odpowiedniości .....	95
7.2.	Wyniki analizy ilościowej .....	97
7.2.1.	Wyniki dla obrazów segmentowanych metodą BGOM .....	98
7.2.2.	Wyniki dla obrazów filtrowanych metodą SSSB .....	102
7.2.3.	Wyniki dla obrazów filtrowanych metodą BLG .....	105
7.3.	Wrażliwość algorytmu na zmianę parametrów .....	108
<b>8.</b>	<b>Badania cech znaczących .....</b>	<b>114</b>
8.1.	Analiza cech w przestrzeni <i>RGB</i> .....	116
8.2.	Analiza cech w przestrzeni <i>HSV</i> .....	121
8.3.	Analiza cech w przestrzeni <i>Lab</i> .....	126
8.4.	Analiza cech w przestrzeni Hough'a .....	130
8.5.	Analiza cech w przestrzeni Canny'ego .....	133
8.6.	Podsumowanie analizy cech .....	139
<b>9.</b>	<b>Dyskusja .....</b>	<b>148</b>
9.1.	Czas parowania <i>t</i> dla różnych obrazów .....	150
9.2.	Skuteczność parowania .....	152

9.3. Liczba błędnych sparowań .....	155
9.4. Podsumowanie dyskusji .....	157
<b>10. Podsumowanie .....</b>	<b>158</b>
<b>Bibliografia.....</b>	<b>161</b>

## 1. CEL I ZAKRES PRACY

Celem pracy jest stworzenie metody pozwalającej na znalezienie odpowiedniości punktów charakterystycznych na obrazach stereowizyjnych. Zaproponowana metoda zostanie wyprofilowana do zastosowań medycznych. Poszukiwanie odpowiedniości odbywać się będzie w dziedzinie obrazów binarnych, co wpłynie na szybkość działania metody. Osiągnięcie celu pracy wymaga spełnienia kilku celów cząstkowych.

1. Opracowanie metody segmentacji obrazów medycznych. Proces segmentacji wyekstrahuje nie tylko pożądane elementy anatomiczne, ale i usunie obszary homogeniczne, które sprawiają dużo trudności podczas procesu poszukiwania odpowiedniości (z powodu braku elementów, które można by porównywać).
2. Opracowanie metody poszukiwania odpowiedniości punktów charakterystycznych. Metoda będzie funkcjonowała w dziedzinie obrazów binarnych z wyraźnie uwidocznionymi elementami np. w postaci krawędzi pionowych, bądź poziomych. Metoda zakłada niewielkie zagęszczenie elementów, wcześniej poprawnie wysegmentowanych.
3. Ocena dokładności powstałej w wyniku poszukiwania odpowiedniości macierzy dysparycji. Przyjęte zostaną trzy kryteria powszechnie stosowane w literaturze. Bazują one na posiadaniu wzorcowych wartości do których porównuje się otrzymane wyniki.
4. Ocena wpływu doboru cech obrazu na jakość wyników poszukiwania odpowiedniości. Bazując na wcześniejszych badaniach autora niniejszej rozprawy, dokonany zostanie wybór cech znaczących, które mają wpływ na wyniki poszukiwania odpowiedniości. Badanie wpływu cech będzie stanowić znaczącą część niniejszej rozprawy.
5. Ostateczna optymalizacja algorytmów do pracy na obrazach naczyń podskórnych. Detekcja naczyń zostanie przeprowadzona z pomocą iluminatora naczyniowego, który dokonuje projekcji obrazu naczyń na skórze pacjenta. Zarejestrowany obraz zawiera oprócz elementów anatomii pacjentów wirtualny obraz naczyń.

Wynikiem działania metody będzie macierz dysparycji, która jest podstawą do późniejszego stworzenia przestrzennego modelu w postaci chmury punktów osi naczyń podskórnych. Przestrzenny model jest podstawą do określenia położenia naczyń względem innych elementów sceny, co umożliwi przeprowadzenie zabiegów medycznych, na przykład

automatycznej iniekcji i infuzji dożylniej, pobierania krwi oraz skleroterapii. Przyczyni się to również do rozwoju systemów nawigacji operacyjnej oraz stacji iniekcji, takich jak bloodbot.

W rozdziale „Wprowadzenie” przedstawiony zostanie proces rekonstrukcji przestrzennej opartej o stereowizję. Omówione zostaną powszechnie stosowane i uznane metody poszukiwania odpowiedniości punktów charakterystycznych na obrazach stereowizyjnych. Przedstawione zostaną atrybuty obrazu, prymitywy oraz metody ich pozyskiwania, które mogą być przydatne z punktu widzenia późniejszego wyboru cech znaczących.

W rozdziale „Materiał badawczy” przedstawione zostaną obrazy, które będą pełnić rolę obrazów testowych podczas wszystkich przeprowadzanych badań. Przedstawiony zostanie również problem pozyskiwania i powstawania dysparycji wzorcowych.

W rozdziale „Opis metody znajdowania odpowiedniości punktów” przedstawione zostaną proponowane metody segmentacji oraz poszukiwania odpowiedniości punktów charakterystycznych, które powstały w wyniku prowadzonych badań. Dla wszystkich metod poszukiwania odpowiedniości zostanie przeprowadzona analiza złożoności obliczeniowej. Przedstawiony zostanie problem cech charakterystycznych oraz wyboru cech znaczących ze względu na proces poszukiwania odpowiedniości.

W rozdziale „Testowanie opracowanego systemu” przedstawione zostaną wyniki działania zaproponowanych metod znajdowania odpowiedniości punktów na obrazach stereowizyjnych w oparciu o dysparycję wzorcowe oraz wpływ wrażliwości algorytmów na zmianę parametrów.

W rozdziale „Badania cech znaczących” zostaną przedstawione wyniki badań nad wpływem wybranych znaczących cech na efektywność poszukiwania odpowiedniości. Zostanie podjęta próba analizy i oceny wyników pod kątem dalszych badań nad odpowiednością punktów charakterystycznych.

W „Dyskusji” przedstawione zostanie porównanie znanych powszechnie stosowanych metod znajdowania odpowiedniości punktów z nowymi, zaproponowanymi w niniejszej rozprawie doktorskiej metodami. Zaprezentowane zostaną wyniki badań świadczące o przewadze zaproponowanej nowej metody nad znanymi metodami w dziedzinie zobrażeń naczyń podskórnych.

W rozdziale „Podsumowanie” sformułowane zostaną wnioski końcowe oraz uwagi dotyczące możliwych kierunków dalszych prac.



## 2. TEZY

W związku z realizowanym tematem rozprawy zaproponowano dwie tezy pracy:

1. Dedykowane metody znajdowania odpowiedniości punktów charakterystycznych na obrazach stereowizyjnych naczyń krwionośnych pozwalają na polepszenie jakości otrzymywanych wyników odpowiedniości.
2. Wprowadzenie nowych cech punktów charakterystycznych na obrazach stereowizyjnych pozwala na podniesienie skuteczności działania metod znajdowania odpowiedniości.

Tezy te wynikają ze wstępnych badań autora dotyczących odpowiedniości punktów oraz założeń doboru odpowiednich cech punktów charakterystycznych. Cechy te zostaną w dalszej pracy wyselekcjonowane i przetestowane w praktyce na obrazach.

Zajmując się w niniejszej pracy metodami poszukiwania odpowiedniości punktów na obrazach stereowizyjnych, należy pamiętać, iż jest to jeden z etapów rekonstrukcji przestrzennej metodą stereowizyjną, a efektem końcowym będzie model przestrzenny. Poszukiwanie odpowiedniości jest kluczowym etapem opracowania modelu przestrzennego. Opracowane metody mogą mieć szerokie zastosowanie, również poza dziedziną medycyny, wszędzie tam, gdzie zachodzi potrzeba opracowania modelu przestrzennego poziomo lub pionowo zorientowanych nieregularnych liniowych obiektów.

Opracowane metody pozwolą na:

- wysegmentowanie krawędzi obiektów i znajdowanie odpowiedniości punktów je definiujących,
- wykluczenie problematycznych powierzchni homogenicznych na etapie segmentacji,
- szybkie znajdowanie odpowiedniości na obrazach binarnych.

Dzięki przedstawionym powyżej cechom, proponowane metody poprawią wyniki odpowiedniości w konkretnych zastosowaniach, w stosunku do powszechnie obecnie stosowanych i znanych metod. Wyniki działania opracowanych nowych metod zostaną porównane ze znanymi metodami poszukiwania odpowiedniości, gdzie wykazana zostanie niższa skuteczność istniejących metod w znajdowaniu odpowiedniości punktów na

wysegmentowanych nieregularnych liniowych elementach znajdujących się na obrazach binarnych.

### 3. PODSTAWOWE DEFINICJE I OZNACZENIA STOSOWANE W PRACY

Poniżej zostały przedstawione **uproszczone definicje** związane z obraną tematyką pracy w kolejności występowania w tekście. Dodatkowo przedstawiono zestawienie oznaczeń stosowanych w pracy.

#### 3.1. Definicje

<i>Pseudokod</i>	Wszystkie algorytmy zostały zaimplementowane w systemie MATLAB® ver. 8.6 R15b. Choć ogólna składnia jest zgodna z C++ to występuje w nim wiele optymalizacji charakterystycznych tylko dla środowiska Matlab. Wymusiło to prezentacje algorytmów w pseudokodzie, dla lepszego zrozumienia kodu, bez wprowadzania elementów języka Matlab.
<i>Funkcje</i>	Opierając się na notacji węgierskiej, nazwy funkcji zaczynają się od małej litery, są zapisane kursywą, a nazwy odzwierciedlają przeznaczenie funkcji np. <i>liczPierwiastek</i> .
<i>Terminy obcojęzyczne</i>	Tłumaczenia prezentowane są kursywą w nawiasach okrągłych np. (ang. <i>matching</i> ).
$m, n$	Odpowiednio wiersz i kolumna macierzy obrazu, iterowane od 1 do odpowiednio $M$ i $N$ .
$O$	Oznacza ogólnie obraz, a dalsze rozróżnienie funkcjonalne następuje przez oznaczenie w dolnym indeksie.
$(L), (P)$	Odpowiednio oznacza lewy i prawy obraz, lub elementy należące do tych obrazów.
$O(m, n)$	Punkt obrazu $o$ znajdujący się w wierszu $m$ i kolumnie $n$ .
$M \times N$	Rozdzielczość obrazu definiowana jako liczba wierszy $m$ (wysokość obrazu) na $n$ kolumn (szerokość obrazu).
[m]	Jednostki podawane są w nawiasach kwadratowych, by jednoznacznie odróżnić je od składników wzorów.
[p]	Stosowane w niniejszej rozprawie oznaczenie jednostki piksel.
1D, 2D, 3D	Akronimy pochodzące z języka angielskiego (ang. <i>one-dimensional</i> , <i>two-dimensional</i> , <i>three-dimensional</i> ) oznaczające odpowiednio przestrzeń

	euklidesową jednowymiarową, dwuwymiarową i trójwymiarową, zarówno w ujęciu liniowym, jak i dyskretnym.
Obraz binarny	Obraz, którego wartość piksela jest przechowywana w jednym bicie danych.
Rekonstrukcja 3D	Proces wiernego odtworzenia kształtu i wyglądu obiektów świata rzeczywistego, zazwyczaj w postaci numerycznej.
Stereowizja	Metoda pasywna rekonstrukcji 3D pozwalającą na określenie współrzędnych przestrzennych sceny w oparciu o jednoczesne podwójne zobrazowanie tego samego obszaru sceny, wykonane z różnych geometrycznie punktów przestrzeni urządzeniami obrazującymi o znanych parametrach.
Stereopara	Obraz lewy i prawy powstały w wyniku obrazowania stereowizyjnego.
Urządzenie obrazujące	Ogólne określenie dla aparatów fotograficznych, kamer video oraz kamer internetowych, pozwalających na cyfrową rejestrację obrazów 2D.
Położenie kardynalne	Układ dwóch urządzeń obrazujących z osiami optycznymi, znajdującymi się na tej samej wysokości, wzajemnie równoległymi i prostopadłymi do wektora translacji, pozwalający na rejestrowanie obrazów zrektyfikowanych.
Głowica stereowizyjna	Układ dwóch urządzeń obrazujących ze środkami rzutów oddalonymi od siebie o wektor translacji, zazwyczaj rozmieszczonych w położeniu kardynalnym.
Aberracja optyczna	Wada soczewki lub układu soczewek, obniżająca przydatność obrazów do celów pomiarowych.
Kalibracja urządzenia obrazującego	Proces w wyniku którego poznajemy geometrię niemetrycznych urządzeń obrazujących (elementy orientacji wewnętrznej) oraz zespołu tych urządzeń (elementy orientacji zewnętrznej). Kalibracja jest konieczna do przeprowadzenia rektyfikacji obrazów oraz triangulacji.
Rektyfikacja obrazu	Doprowadzenie płaszczyzn zobrazowań tworzących stereoparę do koplanarności oraz odpowiadających sobie linii rdzennych do kolinearności.
Geometria płaszczyzn rdzennych	Geometria obrazowań stereowizyjnych, nazywana również z ang. geometrią epipolarną, będąca podstawą dla procesu rektyfikacji obrazów i poszukiwania odpowiedniości wzdłuż promieni rdzennych (linii epipolarnych).
Poszukiwanie odpowiedniości	Znajdowanie odpowiadających sobie elementów na wielokrotnych zobrazowaniach tej samej sceny. Poszukiwanym elementem może być luminancja lub barwa piksela (ewentualnie z otoczeniem zadany oknem) lub prymityw.

Prymityw	Element obrazu, którego wyodrębnienie jest możliwe stosując odpowiednie techniki (np. gradient, laplasjan). Przykładem prymitywów są np. punkty charakterystyczne (ang. <i>corners</i> ), krawędzie, skupiska (ang. <i>blobs</i> ) czy poligony.
Punkty homologiczne	Punkty sparowane w wyniku procesu poszukiwania odpowiedniości.
Dysparycja	Dla zrektyfikowanych zobrazowań stereowizyjnych jest to różnica kolumny $n$ lewego punktu homologicznego i kolumny $n$ prawego punktu homologicznego. Zbiorczą informację o dysparycjach dla całego zobrazowania zawiera macierz dysparycji.
Triangulacja	Algorytm wyliczenia współrzędnych przestrzennych dla homologicznych elementów obrazów dwuwymiarowych, bazujący na wiedzy o elementach orientacji wewnętrznej i zewnętrznej urządzenia obrazującego oraz macierzy dysparycji.
Chmura punktów	Najprostszy model sceny przestrzeni przedmiotowej, powstały w wyniku triangulacji, składający się z punktów posiadających przestrzenne współrzędne, zazwyczaj w układzie lokalnym, na którym można dokonywać pomiarów lub poddawać dalszemu modelowaniu.

### 3.2. Oznaczenia

Poniżej przedstawiono wykaz oznaczeń stosowanych w pracy w kolejności alfabetycznej.

$A$	– macierz Harrisa, macierz autokorelacji,
$B$	– wynik splotu obrazu z filtrem Gaussa,
$D$	– macierz dysparycji,
$D_m$	– pośrodkowa macierz dysparycji,
$D_s$	– mediana macierzy dysparycji,
$d_{max}$	– maksymalna długość pomiędzy dwoma segmentami linii,
$d_{min}$	– minimalną długość segmentu linii,
$D_T$	– macierz dysparycji wzorcowej,
$E$	– macierz podstawowa,
$e$	– punkt epipolarny (rdzenny),
$F$	– macierz fundamentalna,
$f$	– ogniskowa,
$G$	– punkt główny obrazu,
$G_n$ i $G_m$	– pochodna cząstkowa obrazu odpowiednio po $n$ i $m$ ,
$H$	– maska obrazu,
$H_O$	– entropia obrazu,
$i, k$	– dodatkowe rozmiary macierzy,
$I, K$	– liczba elementów dla rozmiaru $i, k$ macierzy,
$Inf$	– wartość symbolizująca nieskończoność,
$J$	– liczba elementów składowych wektora przestrzeni barw,
$j$	– składowe wektora przestrzeni barw,
$K_H$	– detektor Harris’a,
$k_H, k_S, k_V$	– współczynniki cech w przestrzeni barw HSV,
$k_L, k_a, k_b$	– współczynniki cech w przestrzeni barw $L^*a^*b^*$ ,
$K_M$	– detektor Moravec’a
$K_N$	– detektor Noble’a,
$k_R, k_G, k_B$	– współczynniki cech w przestrzeni barw RGB,
$K_{ST}$	– detektor Shi-Tomasi’ego
$l_c$	– liczba Canny’ego,
$l_{max}$	– liczba lokalnych maksimów w przestrzeni Hough’a,

$l_p$	– liczba elementów obrazu, rozumianych jako piksele o wartości 1,
$l_z$	– liczba obrazów,
$M$	– liczba wierszy macierzy obrazu,
$m$	– oznaczenie wiersza macierzy obrazu,
$MI$	– Informacja Wzajemna,
$N$	– liczba kolumn macierzy obrazu,
$n$	– oznaczenie kolumny macierzy obrazu,
$O_b$	– obraz binarny,
$O_{corr}$	– znormalizowana korelacja krzyżowa,
$O_k$	– jeden z obrazów pośrednich (SSSB),
$O_o$	– wstępna macierz odległości lokalnych (MED),
$O_{odl}$	– macierz odległości (MED),
$O_{rk}$	– dodatni gradientu obrazu liczonego po $m$ , (SSSB),
$O_{rw}$	– dodatni gradientu obrazu liczonego po $n$ , (SSSB),
$O_s$	– obrazy z wysegmentowanymi osiami obiektów (BGOM),
$O_f$	– finalna macierz odległości lokalnych (MED),
$O_w$	– jeden z obrazów pośrednich (SSSB),
$O_z$	– obraz zrektyfikowany,
$p$	– dowolna wartość progowa, zazwyczaj próg binaryzacji,
$P$	– punkt w przestrzeni przedmiotowej,
$p_{RGB}$	– wartość progowa binaryzacji dla przestrzeni $RGB$ ,
$p_{HSV}$	– wartość progowa binaryzacji dla przestrzeni $HSV$ ,
$p_{Lab}$	– wartość progowa binaryzacji dla przestrzeni $Lab$ ,
$p_b$	– wartość progowa mediany macierzy dysparycji,
$p_D$	– dopuszczalny próg błędu sparowania,
$p_k$	– wartość progowa (MED- $RGB$ ),
$p_{min}, p_{max}$	– progi histerezy dla operatora Canny’ego,
$p_s$	– wartość progowa (MED-NDD i MED- $RGB$ ),
$q$	– punkt w przestrzeni obrazowej, wyrażony w jednostkach układu obrazu,
$Q$	– punkt w przestrzeni przedmiotowej, wyrażony w jednostkach układu obrazu,
$R$	– macierz rotacji,
$r_1, r_2, r_3$	– współczynniki dystorsji radialnej,
$R_n$	– wektor zero-jedynkowy dla punktu następującego,
$R_n$	– zero-jedynkowy wektor dla elementu następującego (MED- $RGB$ ),

$R_p$	– wektor zero-jedynkowy dla punktu poprzedzającego,
$R_p$	– zero-jedynkowy wektor dla elementu poprzedzającego (MED-RGB),
$S_j$	– średnia składowych barwy $j$ obrazu wymnożonych przez cechy (SSSB),
$S$	– środek rzutów na obrazie,
$S_k$	– suma wartości barwy dla 3 pikseli w kolumnie (SSSB),
$S_n$	– odległość do elementu następującego (MED-NDD),
$S_p$	– odległość do elementu poprzedzającego (MED-NDD),
$S_w$	– suma wartości barwy dla 3 pikseli w wierszu (SSSB),
$T$	– wektor translacji,
$U$	– wektor dystorsji,
$t_1, t_2$	– współczynniki dystorsji tangencjalnej,
$u(n, m)$	– funkcja będąca miarą paralaksy,
$W$	– macierz elementów orientacji wewnętrznej,
$w(u, v)$	– okno, wycinek obrazu o rozmiarach $u$ na $v$ ,
$X, Y, Z$	– współrzędne docelowe w danym układzie współrzędnych,
$Z$	– funkcjonal energii,
$\Delta_n$	– różnica elementów następujących na obu obrazach,
$\Delta_p$	– różnica elementów poprzedzających na obu obrazach,
$\delta_D, \delta_{Dp}$	– procentowa liczba błędnych sparowań,
$\delta_r, \delta_{rp}$	– skuteczność parowania obrazów,
$\varepsilon$	– stała bliska zeru,
$\varepsilon_c$	– otoczenie liczby Canny’ego,
$\kappa$	– parametr w detektorze Harrisa,
$\lambda$	– wartość własna macierzy,
$\rho, \theta$	– współrzędne w parametrycznej przestrzeni Hough’a,
$\sigma$	– odchylenie standardowe filtra Gaussowskiego.



## 4. WPROWADZENIE

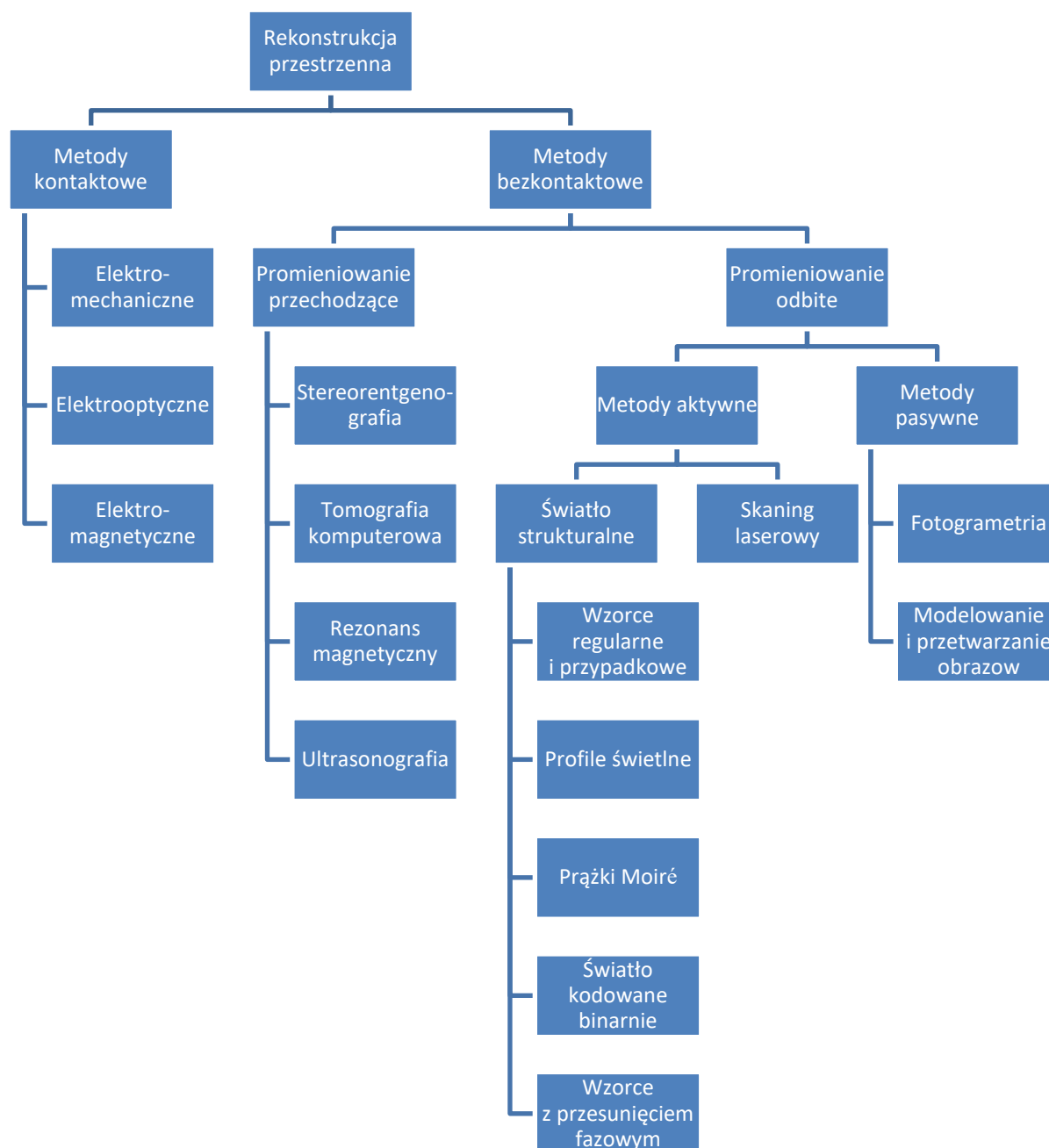
Automatyzacja jest już widoczna w każdej dziedzinie naszego życia. Pozwala ona między innymi zwiększyć precyzję i niezawodność usług czy produktów. Wraz z rozwojem techniki mikroprocesorowej, dzięki komputerom automatyzacja wkroczyła w nową erę, przykładowo dając nadzieje pacjentom na protezy nowej generacji [1, 2]. Następnym kamieniem milowym automatyzacji było powstanie i szybki rozwój analizy obrazu (ang. *computer vision*).

Dzięki osiągnięciom analizy obrazu, wiele dziedzin otrzymało nowe zaawansowane narzędzia, pozwalające podnieść ich efektywność. Do tych dziedzin zaliczyć należy rekonstrukcję 3D, która pozwala uzyskiwać wirtualne modele, precyzyjnie przystające do rzeczywistych obiektów, celem np. dalszej ich analizy.

Rekonstrukcja 3D posiada wiele zastosowań, z których niektóre zostaną poniżej wymienione.

1. Wirtualizacja medyczna. Pozwala na przeniesienie sali operacyjnej do wirtualnego świata, by następnie przeprowadzać zdalnie zabiegi, czy asystować w operacjach wymagających szczególnej precyzji działań ze strony chirurga.
2. Diagnostyka medyczna wad postawy, w tym szczególnie skrzywienia kręgosłupa oraz diagnostyka chorób metabolicznych. Znacznie ułatwia i przyspiesza przeprowadzanie programów badań przesiewowych oraz pozyskania wyników i ich prezentacji.
3. Chirurgia szczękowo-czaszkowa oraz plastyczna. Pozwala na szybką analizę kształtu kości, odtworzenie brakujących fragmentów, czy symulację wyników operacji.
4. Systemy treningu i analizy osiągnięć w sporcie. Śledzenie trajektorii ruchów stawów sportowców osiągających mistrzowskie wyniki są nakładane na ruchy innych zawodników i umożliwiają wprowadzanie we właściwe trajektorie.
5. Kontrola jakości, szczególnie przydatna w fabrykach produkujących elementy wymagające wysokiej dokładności wykonania.
6. Opracowania inżynierskie z zakresu geodezji i fotogrametrii oraz architektury. Budowa modelu 3D terenu i budynków stanowiła jedno z pierwszych zastosowań rekonstrukcji 3D.
7. Konserwacja i rekonstrukcja zabytków. Utworzony model 3D, można zarchiwizować i w dowolnym czasie dokonać rekonstrukcji zniszczonych zabytków.

8. Przemysł odzieżowy. Po stworzeniu awatara można szybko dokonywać wirtualnych przymiarek w sklepie jak i na odległość.



Rys. 4.1 Podział metod rekonstrukcji 3D ze względu na inwazyjność.

9. Animacja i przemysł rozrywkowy. Podobnie jak wyżej spersonalizowany awatar dostarcza nowych doznań graczom.

Na rys. 4.1 pokazano podział metod rekonstrukcji przestrzennej przyjmując jako kryterium podziału inwazyjność metody [3].

Wśród metod pasywnych umiejscowione są metody oparte o analizę obrazu, czyli modelowanie i przetwarzanie obrazów. Szczególnie interesujący jest potencjał aplikacyjny tych metod w dziedzinie medycyny.

#### **4.1. Rekonstrukcja 3D w medycynie**

W dziedzinie medycyny obserwuje się szczególne zainteresowanie i zapotrzebowanie automatyzacją procedur medycznych [4]. Głównym powodem tego zainteresowania są coraz wyższe wymagania dotyczące precyzji wykonywanych czynności, przekładające się na podwyższoną skuteczność zabiegów oraz minimalizację działań niepożądanych, a także przyspieszenie rekonwalescencji [5]. Już w latach 90 poprzedniego wieku prowadzone były przez fotogrametrów i informatyków prace nad systemami rekonstrukcji 3D twarzy [6]. w początkach wieku wdrożono fotogrametryczny system do pomiaru wad postawy [7] i systemy tworzące pełny model ludzkiego ciała [8].

Wirtualizacja medyczna dzięki zastosowaniu rekonstrukcji 3D umożliwia stworzenie modelu pacjenta i narzędzi chirurgicznych. Do wspólnego układu współrzędnych są rejestrowane wszystkie zobrazowania diagnostyczne, takie jak między innymi CT, czy MRI, zarówno wykonywane wcześniej, jak i na bieżąco podczas operacji. Połączone modele pacjenta, narzędzi chirurgicznych oraz zobrazowań diagnostycznych można podczas operacji obserwować na sterylnych wielkogabarytowych wyświetlaczach. Na bieżąco wyświetlane są również informacje nawigacyjne dla narzędzi chirurgicznych, z których korzysta chirurg.

Opisany powyżej system nazywany został systemem nawigacji śródoperacyjnej i stworzony został pierwotnie na potrzeby chirurgii czaszkowej [9]. Od lat dziewięćdziesiątych ubiegłego wieku zanotowały one szybki rozwój i upowszechnienie się na całym świecie. Dziś systemy nawigacji wykorzystuje się między innymi do operacji kręgosłupa [10, 11], endoprotezoplastyki kolana [12, 13], alloplastyki stawu biodrowego [14], a nawet w operacjach laryngologicznych zatok przynosowych [15, 16].

Funkcjonujące dziś systemy nawigacji bazują na urządzeniach obrazujących w zakresie podczerwieni. Przyrządy chirurgiczne, zarówno w systemach pasywnych, jak i aktywnych, zaopatrzone są w specjalne markery lub diody emitujące podczerwień. Budowa modelu przestrzennego polega na rejestracji charakterystycznych punktów na powierzchni ciała, by następnie wpasować w te punkty wcześniejsze zobrazowania MRI i CT, jak również na bieżąco wykonane zobrazowania mobilnymi wersjami tomografów.

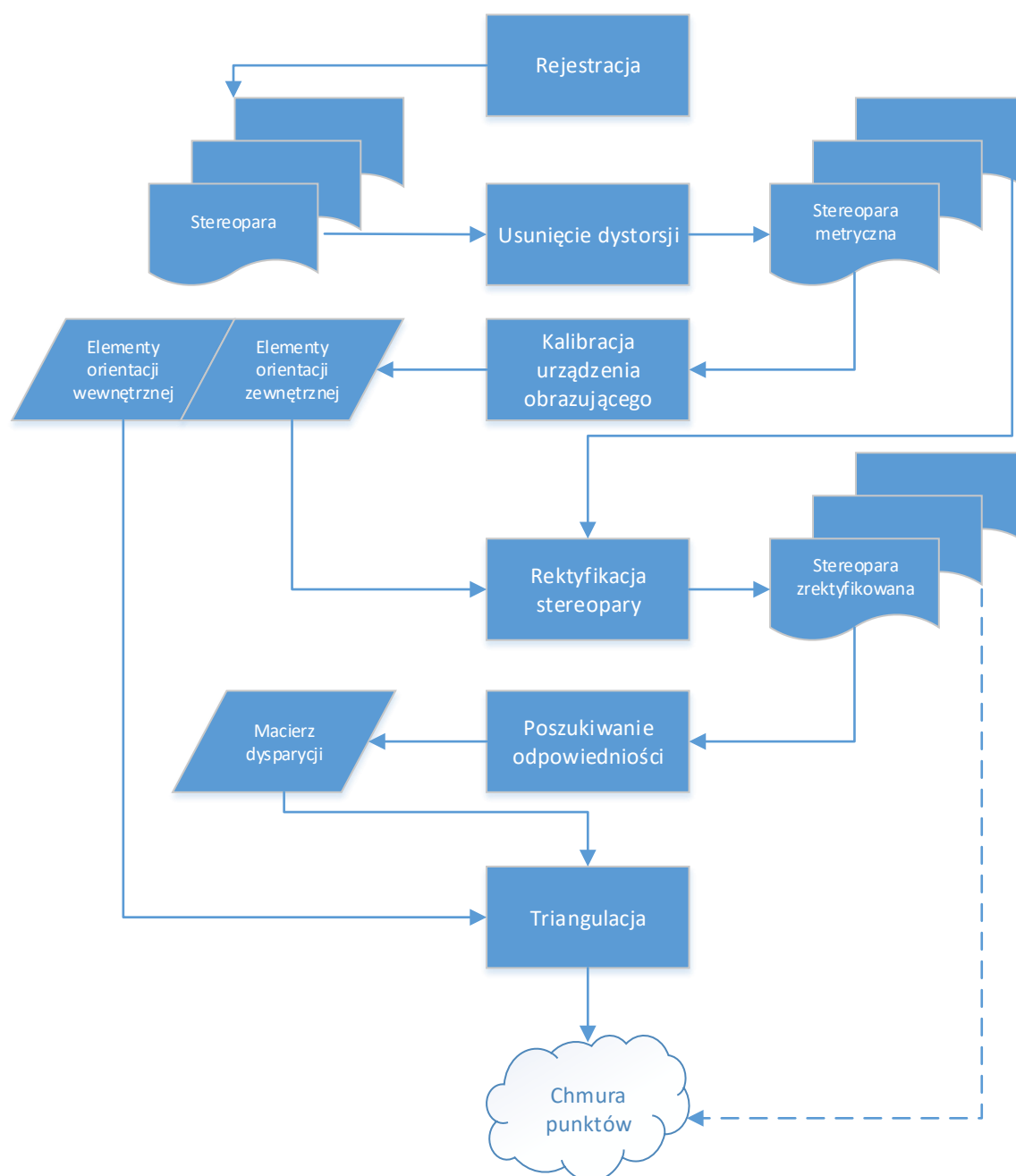
Naturalnym uzupełnieniem systemów nawigacji śródoperacyjnej, mogą być stacje iniekcji, w postaci robotycznego ramienia z igłą, zintegrowanego z skanerem naczyń podskórnych oraz dedykowanym systemem obrazującym. Możliwe jest również wykorzystanie, natywnych dla systemów nawigacji śródoperacyjnej, urządzeń obrazujących w podczerwieni. System taki może mieć duże znaczenie w przypadku pacjentów pediatrycznych, u których pobieranie krwi żyłnej stanowi problem z powodu niewielkiej średnicy naczyń oraz trudności w unieruchomieniu [17].

Skonstruowana przez Imperial College London stacja iniekcji nazwana bloodbot [18], bazuje na badaniu siły nacisku głowicy z igłą i oporu stawianego przez skórę. Zaproponowane w niniejszej rozprawie doktorskiej algorytmy parowania obrazów binarnych, stwarzają nowe źródło informacji o położeniu naczyń podskórnych. Dzięki temu system korzystając z dwóch niezależnych źródeł informacji o położeniu naczyń staje się bezpieczniejszy w użytkowaniu.

## **4.2. Stereowizja, uzasadnienie wyboru metody**

W niniejszej rozprawie doktorskiej rekonstrukcję 3D przeprowadza się metodą stereowizyjną. Zawdzięcza ona swoje powstanie fotogrametrii, gdzie długo była wykorzystywana w swojej wersji analogowej. o wyborze stereowizji jako metody rekonstrukcji 3D zadecydowały jej zalety:

- relatywnie niewielki koszt głowicy stereowizyjnej,
- obiektywizm pomiaru, uzyskiwany poprzez zdalne pozyskiwanie danych i brak kontaktu z mierzonym obiektem, co daje brak wpływu na wyniki pomiaru,
- całkowicie zdalny, bezdotykowy pomiar,
- dzięki wykorzystaniu światła odbitego metoda jest w 100 % bezpieczna dla zdrowia,
- czas pomiaru rzędu dziesiątych lub setnych części sekundy,
- możliwość pomiaru dowolnej liczby punktów na obrazach, które traktuje się jako zarchiwizowane zbiory danych, które mogą być wykorzystywane dowolną liczbę razy,
- wysoka dokładność pomiaru, dla tzw. fotogrametrii bliskiego zasięgu (odległość obrazowania poniżej 300 m) rzędu 1/10 000 tej odległości,
- wielka różnorodność mierzonych obiektów i zjawisk na nich zachodzących uzyskana przez możliwość rejestracji w różnych zakresach długości fal,



Rys. 4.2 Poszczególne etapy rekonstrukcji przestrzennej metodą stereowizyjną.

- rejestracja procesów dynamicznych,
- przy obecnym stanie rozwoju informatyki możliwość całkowitej automatyzacji pomiaru.

Proces tworzenia modelu 3D w stereowizji na przestrzeni ostatnich dwóch dekad ulegał ciągłym modyfikacjom w zależności od postępu metod i technik optymalizacyjnych. w niniejsze rozprawie doktorskiej przyjęto model stereowizji przedstawiony na rys. 4.2 bazujący na obrazach zrektyfikowanych.

W następnych podrozdziałach, o nazwach korespondujących z rys. 4.2, przybliżone zostaną poszczególne etapy opracowania modelu przestrzennego stosowanego w stereowizji, zaczynając od zagadnienia rejestracji obrazów, poprzez problem usunięcia aberracji, kalibracji, rektyfikacji, poszukiwania odpowiedniości, a na triangulacji współrzędnych chmury punktów skończywszy.

### 4.3. Rejestracja

Stereowizja, tradycyjnie bazuje na dwóch obrazach, tzw. stereoparze. Rejestracja ich może być wykonana za pomocą:

- a) specjalnych metrycznych urządzeń obrazujących, pozbawionych aberracji i o znanej geometrii,
- b) niemetrycznych urządzeń obrazujących, o nieznanych wartościach aberracji i geometrii.

Rejestracja może być wykonana w następujący sposób:

- c) dwuobrazowym urządzeniem obrazującym, które rejestruje jednocześnie dwa obrazy z odpowiednio ustawioną bazą,
- d) dwoma jednoobrazowymi urządzeniami obrazującymi, oddalonymi od siebie o wartość bazy,
- e) jednoobrazowym urządzeniem obrazującym pierwszy obraz, by następnie przemieścić się na odległość bazy i dokonać rejestracji drugiego obrazu.

W zależności od urządzenia obrazującego i sposobu rejestracji proces rekonstrukcji 3D będzie się zaczynał na innym etapie opracowania zgodnie z modelem stereowizji przedstawionym na rys. 4.2. Ze względu na zastosowane urządzenie obrazujące oraz sposób rejestracji wyróżniamy trzy przypadki:

- 1) a+c – rekonstrukcję rozpoczniemy od rektyfikacji stereopary.
- 2) a+d, a+e – rekonstrukcję rozpoczniemy od określenia elementów orientacji zewnętrznej.
- 3) b+c, b+d, b+e – rekonstrukcję rozpoczniemy od usunięcia dystorsji.

W następnym podrozdziale przedstawione zostanie zagadnienie usunięcia błędów optycznych, którymi są obciążone obrazy.

#### 4.4. Usunięcie dystorsji

Stereowizja opiera się na prostym modelu kamery perspektywicznej, zwanym również modelem kamery otworkowej (ang. *pinhole camera model*). Ponieważ jest to prosty model oparty o geometrię układu, nie obejmuje on różnorodnych zniekształceń powodowanych przez układy optyczne. w związku z tym obrazy muszą być pozbawione wszelkich aberracji, czyli błędów optycznych, a w szczególności dystorsji. Istnieje wiele rodzajów aberracji [19], np.: aberracja chromatyczna i sferyczna, koma, dystorsja, czy astygmatyzm. Wszystkie te aberracje można wyeliminować na etapie projektowania i produkcji obiektywu optycznego. Wyjątkiem jest dystorsja, której nie da się usunąć przez odpowiednią konstrukcję obiektywu. Wpływ dystorsji na geometrię obrazu pokazany został na rys. 4.3. Dystorsję dzielimy na radialną i tangencjalną [20]. Dystorsja radialna jest związana z kulistym kształtem soczewek, z których jest zbudowany obiektyw. Wyeliminowanie tego błędu na etapie produkcji wymagało by np. wykonania soczewek parabolicznych lub umieszczenia przysłony dokładnie w geometrycznym środku rzutu.

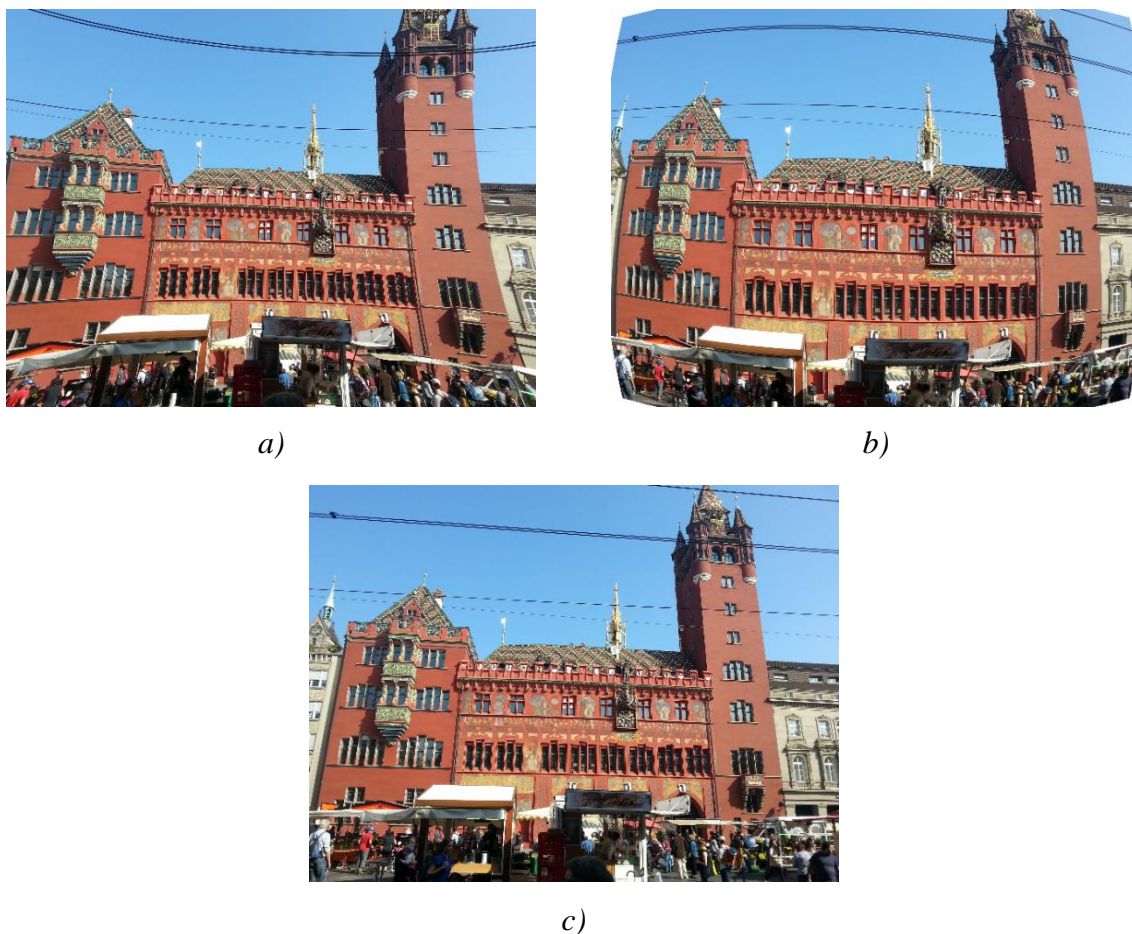
Dystorsja tangencjalna jest związana z nierównoległym umieszczeniem zespołu soczewek względem siebie nawzajem, jak i względem płaszczyzny obrazowej.

W pracy [20] podana jest formuła pozwalająca na jednoczesne usunięcie obu rodzajów dystorsji wprowadzając pięcioelementowy wektor dystorsji  $U$  zdefiniowany jako

$$U = (r_1, r_2, r_3, t_1, t_2) \quad (4.1)$$

gdzie  $r_1, r_2$  i  $r_3$  – współczynniki dystorsji radialnej,  
 $t_1, t_2$  – współczynniki dystorsji tangencjalnej.

Otrzymane w ten sposób 5 parametrów tworzy wektor dystorsji. Korzystając z przekształcenia homograficznego, dysponując zobrazowaniem płaskiego elementu o znanych współrzędnych przedmiotowych np. szachownicy, jesteśmy w stanie znaleźć współczynniki dystorsji rozwiązując układ równań z pięcioma niewiadomymi.



Rys. 4.3 Zniekształcenia geometryczne obrazu wywołane dystorsją: *a)* dystorsja dodatnia (poduszkowata); *b)* dystorsja ujemna (baryłkowata); *c)* obraz z usuniętą dystorsją.

Dysponując współrzędnymi czterech skrajnych narożników, otrzymuje się  $2 \times 4$  równań, co umożliwia wyliczenie pięcioelementowego wektora dystorsji. Znając wektor dystorsji eliminuje się zniekształcenie obrazu zgodnie z [21]. Dopiero po usunięciu dystorsji z obrazów, można przystąpić do kalibracji urządzeń obrazujących, celem poznania ich geometrii wewnętrznej oraz wzajemnego położenia, co zostanie przedstawione w następnych podrozdziałach.

#### 4.5. Elementy orientacji wewnętrznej

Przekształcenie rzutowe odwzorowujące przestrzeń przedmiotową na przestrzeń obrazową można zapisać jako iloczyn macierzowy [22], w postaci równania



$$q = W \cdot Q \quad (4.2)$$

gdzie  $q$  – punkt w przestrzeni obrazowej,

$Q$  – punkt w przestrzeni przedmiotowej,

$W$  – macierz elementów orientacji wewnętrznej zdefiniowana jako

$$W = \begin{bmatrix} f_n & 0 & G_n \\ 0 & f_m & G_m \\ 0 & 0 & 1 \end{bmatrix}. \quad (4.3)$$

Elementami macierzy  $W$  są elementy orientacji wewnętrznej, czyli ogniskowa  $f$  oraz współrzędne punktu głównego  $G$ . Współrzędne jednorodne w przestrzeni obrazowej  $q$  są wynikiem przekształcenia współrzędnych z przestrzeni przedmiotowej  $Q$  przez macierz orientacji wewnętrznej  $W$ . Znając współrzędne w przestrzeni przedmiotowej i obrazowej, otrzymamy układ równań z 4 niewiadomymi, z których możemy obliczyć elementy orientacji wewnętrznej.

Rozpatrując dalej obiekt w przestrzeni przedmiotowej i obrazowej [23, 24] możemy napisać następującą równość

$$q = R \cdot (Q - T) \quad (4.4)$$

gdzie:

$R$  - jest macierzą rotacji pomiędzy układem współrzędnych na obrazie, a układem współrzędnych w przestrzeni przedmiotowej,

$T$  - to wektor translacji pomiędzy początkiem układu współrzędnych na obrazie, a początkiem układu współrzędnych w przestrzeni przedmiotowej.

Macierze  $R$  i  $T$  nazywamy elementami orientacji zewnętrznej. Ponieważ rotacje tworzą trzy kąty, a translację wektor trzelementowy, to ostatecznie otrzymujemy układ równań z 6 niewiadomymi. Po połączeniu układu równań (4.4) z równaniem (4.2) otrzymujemy układ równań z 10 niewiadomymi dla każdego obrazu.

Proces w wyniku którego obliczane są elementy orientacji wewnętrznej i zewnętrznej, nazywa się kalibracją urządzenia obrazującego. Można go przeprowadzić wieloma metodami, jednak najbardziej jest rozpowszechniona metoda bazująca na zobrazowaniach planszy testowej z naniesioną szachownicą o znanej geometrii. Dysponując 4 najbardziej zewnętrznymi punktami szachownicy, deficyt równań nie pozwoli na obliczenie wszystkich niewiadomych.

Zakładając, iż dysponuje się  $l_z$  obrazami, gdzie każdy obraz ma 4 narożniki, otrzymuje się  $2 \cdot 4 \cdot l_z$  układów równań, pomnożone przez 2, ponieważ w przestrzeni obrazowej dysponuje się dwoma wymiarami.

Wykonując zobrazowania szachownicy należy zadbać, by każde zobrazowanie wykonane było pod innym kątem. Jeśli różnice w zobrazowaniach będą za małe, układ równań nie będzie zbieżny, lub otrzymane rozwiązanie będzie błędne.

Wynika z tego, że dla każdego zobrazowania należy obliczyć 6 elementów orientacji zewnętrznej. Natomiast 4 elementy orientacji wewnętrznej są takie same dla wszystkich obrazów. w konsekwencji powyższego otrzymuje się następujące równanie

$$2 \cdot 4 \cdot l_z \geq 6 \cdot l_z + 4. \quad (4.5)$$

W wyniku otrzymuje się  $l_z \geq 2$ . Biorąc pod uwagę błędy przypadkowe i numeryczną stabilność układu równań oczywista jest potrzeba większej liczby obrazów. Praktyczne doświadczenia autora niniejszej rozprawy pokazują, iż dokładną kalibrację uzyskuje się dla co najmniej 15 zobrazowań.

W konsekwencji powyższego otrzymamy macierz elementów orientacji wewnętrznej dla każdego urządzenia obrazującego. Dysponując znajomością geometrii każdego urządzenia obrazującego z osobna, możemy przystąpić do wyliczenia elementów orientacji zewnętrznej głowicy stereowizyjnej.

#### 4.6. Elementy orientacji zewnętrznej

By obliczyć elementy orientacji zewnętrznej dla dwóch urządzeń obrazujących należy wyrazić rzut punktu  $P$ , znajdującego się w przestrzeni przedmiotowej, co zostało przedstawione na rys. 4.4, na obrazie lewym

$$P^{(L)} = R^{(L)} \cdot P + T^{(L)} \quad (4.6)$$

i prawym

$$P^{(P)} = R^{(P)} \cdot P + T^{(P)} \quad (4.7)$$

gdzie:

$P^{(L)}, P^{(P)}$  współrzędne rzutu punktu  $P$  na płaszczyznę obrazową

$R^{(L)}, R^{(P)}$  macierz obrotu układu przedmiotowego do układu obrazu lewego i odpowiednio do układu obrazu prawego,

$T^{(L)}, T^{(P)}$  wektor translacji układu przedmiotowego do układu obrazu lewego i odpowiednio do układu obrazu prawego.

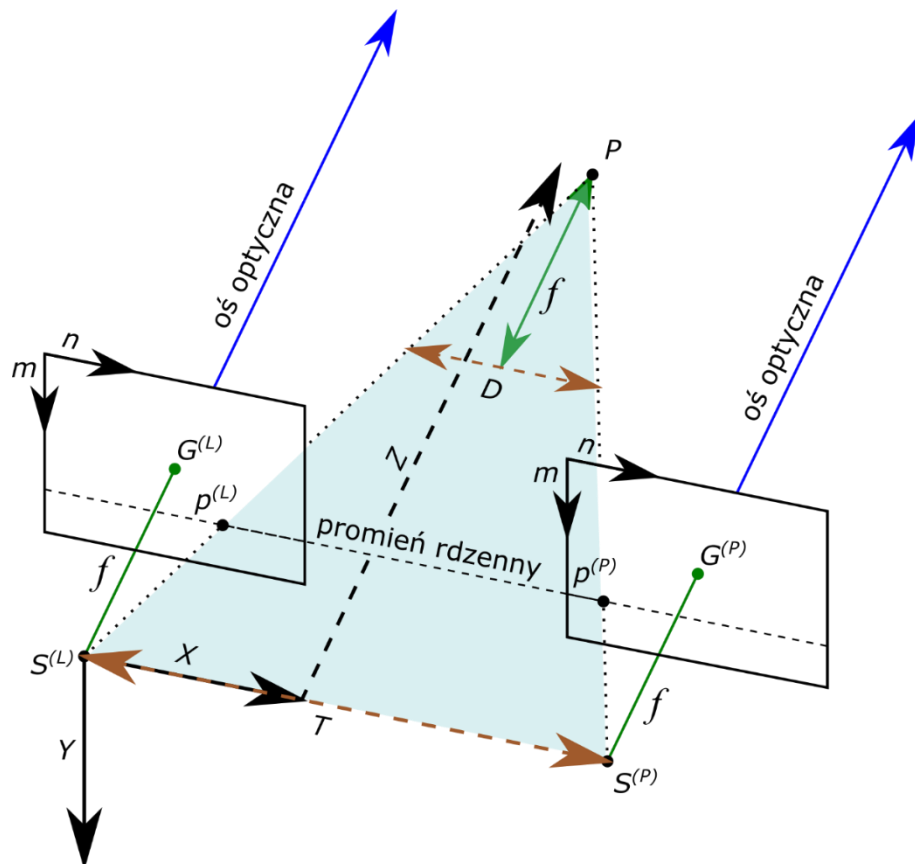
Wielkości te zostały już wcześniej obliczone podczas wyliczania elementów orientacji wewnętrznej. Jednocześnie z zależności geometrycznej współrzędne układu obrazu lewego są funkcją współrzędnych układu obrazu prawego zgodnie ze wzorem

$$P^{(L)} = R^T \cdot (P^{(P)} + T) \quad (4.8)$$

gdzie:

$R$  to szukana przez nas macierz rotacji obrazu prawego w stosunku do lewego,

$T$  wektor translacji obrazu prawego do lewego.



Rys. 4.4. Schemat głowicy stereowizyjnej w ustawieniu kardynalnym:  $P$  punkt w przestrzeni przedmiotowej;  $p^{(L)}, p^{(P)}$  obraz punktu  $P$  na obrazie lewym i prawym;  $X, Y, Z$  układ współrzędnych przedmiotowych;  $T$  baza obrazowania (wektor translacji);  $G^{(L)}, G^{(P)}$  punkt główny na obrazie lewym i prawym;  $S^{(L)}, S^{(P)}$  środek rzutów na lewym i prawym obrazie.

Rozwiązując równania (4.6)(4.7)(4.8) ze względu na  $R$  i  $T$  otrzymujemy

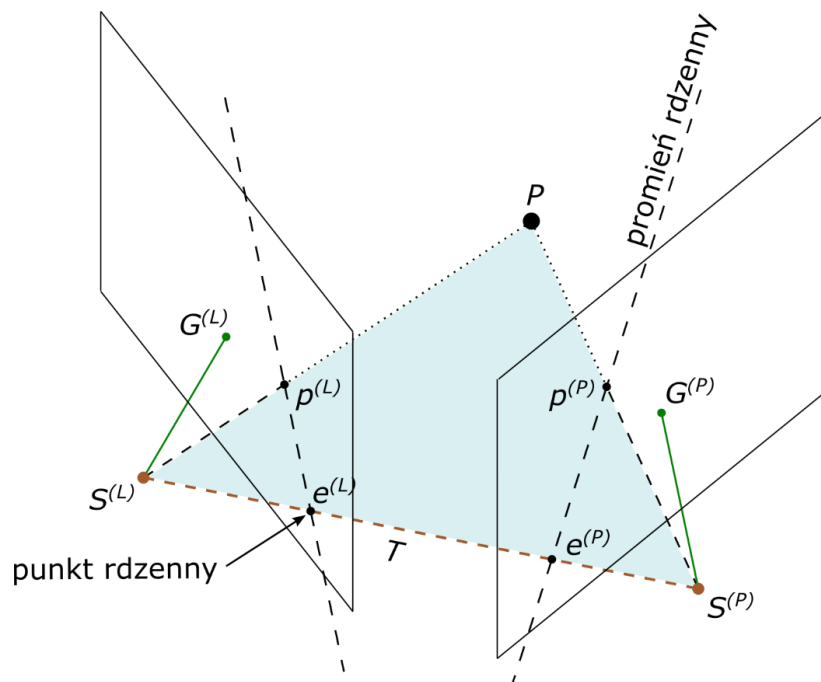
$$\begin{aligned} R &= R^{(P)} \cdot R^{(L)T} \\ T &= T^{(P)} - R \cdot T^{(L)} \end{aligned} \quad (4.9)$$

czyli poszukiwane przez nas elementy orientacji zewnętrznej.

W wyniku kalibracji głowicy stereowizyjnej otrzymuje się elementy orientacji wewnętrznej i zewnętrznej, które umożliwiają przeprowadzenie rektyfikacji obrazów oraz triangulację współrzędnych. w następnym podrozdziale przedstawione zostanie zagadnienie rektyfikacji obrazów tworzących stereoparę.

#### 4.7. Rektyfikacja obrazów

W pracy [25] oraz [26] przedstawiona została szczegółowo geometria epipolarna, która stworzyła teoretyczne podstawy do rektyfikacji obrazów. Wiąże ona ze sobą dwa zobrazowania tej samej sceny, wprowadzając pojęcie płaszczyzny epipolarniej. Płaszczyzna epipolarna jest zdefiniowana przez dany punkt sceny  $P$  oraz środki rzutów obrazu lewego  $S^{(L)}$  i prawego  $S^{(P)}$ , co zostało przedstawione na rys. 4.5 w postaci obszaru wypełnionego turkusową barwą.



Rys. 4.5 Geometria płaszczyzn rdzennych (geometria epipolarna).

Przecina ona płaszczyznę zobrazowania tworząc linie epipolarne. Obrazem prawego środka rzutów na zobrazowaniu lewym jest punkt epipolarny  $e^{(L)}$  i vice-versa dla zobrazowania prawego  $e^{(P)}$ . Należy wyjaśnić, że choć przyjęło się już używać w języku polskim określenia „geometria epipolarna”, to polskimi odpowiednikami używanymi w matematyce są „geometria płaszczyzn rdzennych”, „płaszczyzna rdzenna”, „promień rdzenny” i „punkt rdzenny”.

W geometrii epipolarnej [27] wprowadza się pojęcia macierzy podstawowej  $E$  i macierzy fundamentalnej  $F$ , które łączą punkty lewego obrazu z liniami epipolarnymi na prawym obrazie i vice-versa. Macierz podstawowa  $E$  jest powiązana z obrazami punktu  $P$  na obrazie lewym  $P^{(L)}$  i prawym  $P^{(P)}$  iloczynem skalarnym w postaci

$$P^{(P)T} \cdot E \cdot P^{(L)} = 0. \quad (4.10)$$

Jest ona wyrażona w jednostkach układu przedmiotowego lub jednostkach układu obrazowego, czyli w metrach lub milimetrach. Macierz podstawowa zawiera w sobie elementy orientacji zewnętrznej.

Macierz fundamentalna  $F$  jest powiązana z odpowiadającymi sobie punktami  $q^{(P)}$  i  $q^{(L)}$  w przestrzeni obrazowej iloczynem skalarnym w postaci

$$q^{(P)T} \cdot F \cdot q^{(L)} = 0. \quad (4.11)$$

Jest ona wyrażona w pikselach, dzięki wprowadzeniu do macierzy podstawowej  $E$  macierzy elementów orientacji wewnętrznej  $W$ .

Z geometrii epipolarnej wypływa warunek, który mówi, iż obrazowi punktu  $P^{(L)}$  na zobrazowaniu lewym odpowiada tylko jedna o znanych parametrach linia epipolarna na zobrazowaniu prawym i vice versa. Jest to wynikiem defektu zarówno macierzy  $E$  jak i  $F$ , ponieważ są one rzędu drugiego, przy wymiarach  $3 \times 3$ .

Bazując na warunku epipolarnym parowanie punktów homologicznych, czyli znajdowanie odpowiedniości obrazów, można uprościć do zadania w przestrzeni 1D. z warunku tego wynika oczywisty fakt, że odpowiednik punktu na obrazie lewym leży na dokładnie jednej linii na obrazie prawym. Jeśli dodatkowo założymy, że wszystkie linie epipolarne na obu obrazach są poziome, oraz dodatkowo odpowiadające sobie linie epipolarne na obu obrazach leżą na tym samym poziomie, to znacząco uproszczony zostaje proces poszukiwania odpowiedniości.

Doprowadzenie obu obrazów do takiego położenia nazywamy rektyfikacją. Oczywiście rektyfikację można osiągnąć na dwa sposoby. Już wykonując zobrazowania można ustawić

urządzenia obrazujące w taki sposób, by osie celowe obiektywów były idealnie poziome i równoległe do siebie oraz na tym samym poziomie, a płaszczyzny obrazowe znajdowały się w tej samej pionowej płaszczyźnie. Jest to ustawienie tzw. kardynalne do którego się dąży konstruując głowicę stereowizyjną. Jeśli nie jesteśmy w stanie zapewnić takiego ułożenia urządzeń obrazujących podczas ekspozycji, możemy już wykonane zobrazowania doprowadzić do takiego położenia stosując odpowiednie transformacje.

Dysponując elementami orientacji zewnętrznej głowicy stereowizyjnej procedurę rektyfikacji można wykonać stosując algorytm Bouguet'a [28], który zapewnia minimalizację zniekształceń obrazów, przy jednoczesnej maksymalizacji obszaru wspólnego.

Po doprowadzeniu stereopary do położenia kardynalnego można przystąpić do procedury poszukiwania stereokorespondencji, która zostanie opisana w następnym rozdziale.

#### 4.8. Poszukiwanie odpowiedniości na obrazach

Najbardziej złożonym etapem rekonstrukcji przestrzennej jest poszukiwanie odpowiedniości. Inne stosowane nazwy określające ten etap to parowanie obrazów, pasowanie obrazów, poszukiwanie stereokorespondencji lub po prostu stereokorespondencja (ang. *matching*).

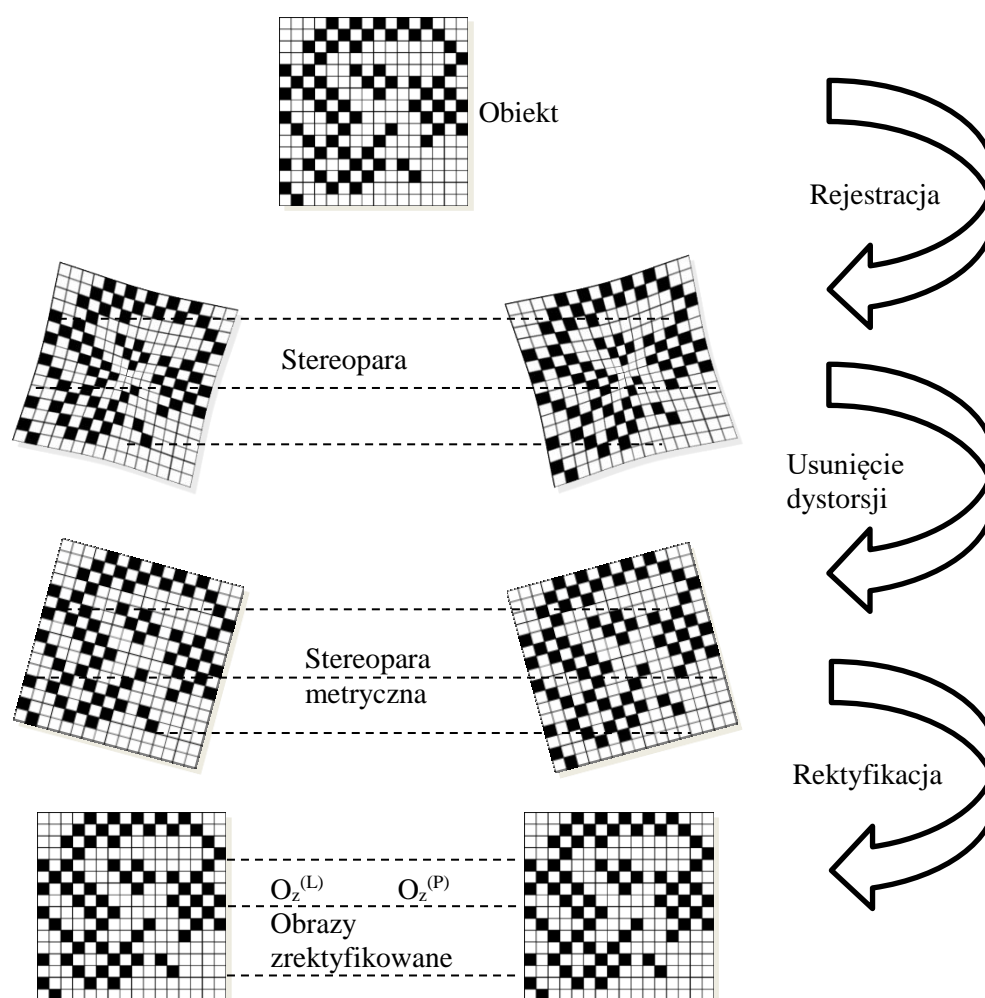
Proces przygotowania obrazów do ich parowania przedstawiony jest na rys. 4.6. Stereokorespondencja nie musi być realizowana w oparciu o geometrię epipolarną. Jednak wcześniejsze zrektyfikowanie obrazów znacząco redukuje złożoność obliczeniową procesu.

Wynikiem parowania obrazów jest macierz dysparycji  $D$ , która zawiera różnicę  $\Delta n$  dla punktów homologicznych, pomiędzy obrazem lewym a prawym.

Parowanie odbywa się w przestrzeni cech przedstawionych na rys. 4.7 i może dotyczyć określonych maską obszarów, w tym pojedynczego piksela (maska równa 1), lub prymitywów, takich jak punkty charakterystyczne, krawędzie, skupiska (ang. *blobs*), poligony [29]. w wyniku parowania obszarów (ang. *ABM - area base matching*, *BM - block matching*) otrzymujemy gęstą macierz dysparycji, podczas gdy wynikiem parowania prymitywów (ang. *feature base matching*) będzie macierz rzadka. Powyższe metody są przykładami metod lokalnych, czyli bazujących na analizie informacji zawartej w bezpośrednio bliskim otoczeniu parowanego elementu.

Powszechnie stosowanym kryterium poprawności parowania elementów jest funkcja kosztu. Za poprawną wartość dysparycji uznaje się tą, której odpowiada minimalna wartość

funkcji kosztu. Zestawienie stosowanych funkcji kosztu z omówieniem można znaleźć w [30]. Do zalet metod lokalnych należy ich prostota i relatywnie niska złożoność obliczeniowa.

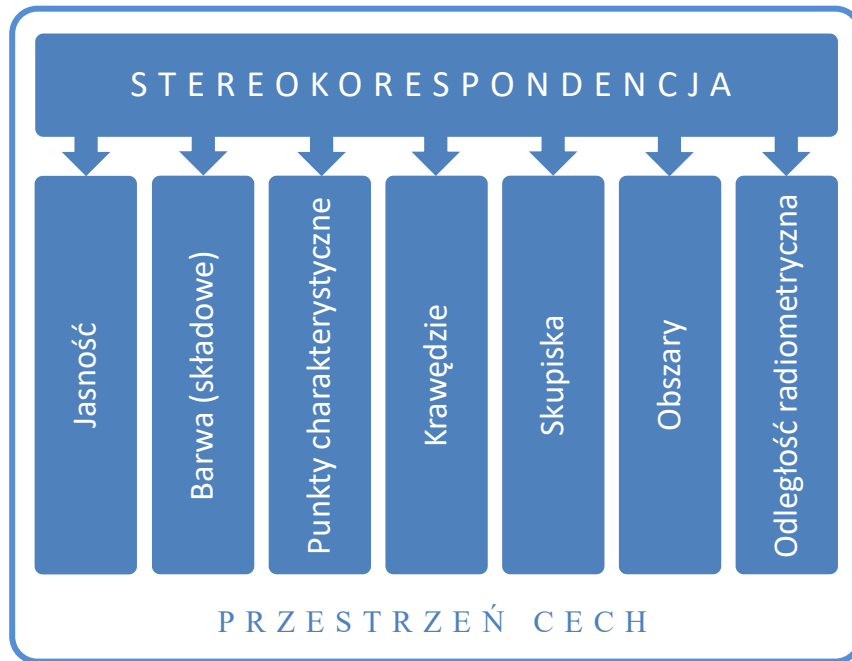


Rys. 4.6 Przygotowanie obrazów do procesu poszukiwania stereokorespondencji.

We wcześniejszych badaniach autora niniejszej rozprawy [31], kiedy opracowywane obrazy nie były poddawane rektyfikacji, należało przed określeniem stereokorespondencji na obu obrazach zlokalizować założone obiekty, a następnie szukać odpowiedniości pomiędzy tymi obiektami. Najczęściej poszukiwało się punktów charakterystycznych, gdzie najwyższą efektywność demonstrowały metody Harrisa [32] i Nobela [33], a następnie poszukiwano pomiędzy nimi odpowiedniości znormalizowaną korelacją krzyżową (ang. *normalised cross correlation* - NCC) [34]. Proces jednak nie zapewniał wystarczającej prędkości do zastosowań w czasie rzeczywistym.

Dzięki rektyfikacji zobrazowań kosztowny obliczeniowo proces detekcji punktów został wyeliminowany. Zamiast tego punkty są kolejno wybierane na obrazie lewym, by następnie na

obrazie prawym rozpocząć poszukiwanie wybraną funkcją podobieństwa odpowiadającego mu punktu na korespondującej linii rdzennej [35], co przedstawiono na rys. 4.8.



Rys. 4.7 Przykładowe cechy podlegające parowaniu.

Poszukiwanie rozpoczyna się na obrazie prawym od pozycji  $n^{(L)}$  – współrzędnej piksela na obrazie lewym, zmierzając w kierunku lewym. Ponieważ obrazy są zrektyfikowane, wartości wierszy  $m$  odpowiadających sobie pikseli są sobie równe, zgodnie z równaniem (4.12).

$$m^{(L)} = m^{(P)} \quad (4.12)$$

Kiedy funkcja podobieństwa odpowie właściwą wartością dla aktualnej pozycji na obrazie prawym  $n^{(P)}$ , co w przypadku NCC oznacza wartość bliska jedności, następuje obliczenie dysparycji w sposób zgodny z równaniem (4.13).

$$D(m, n) = n^{(L)} - n^{(P)} \quad (4.13)$$

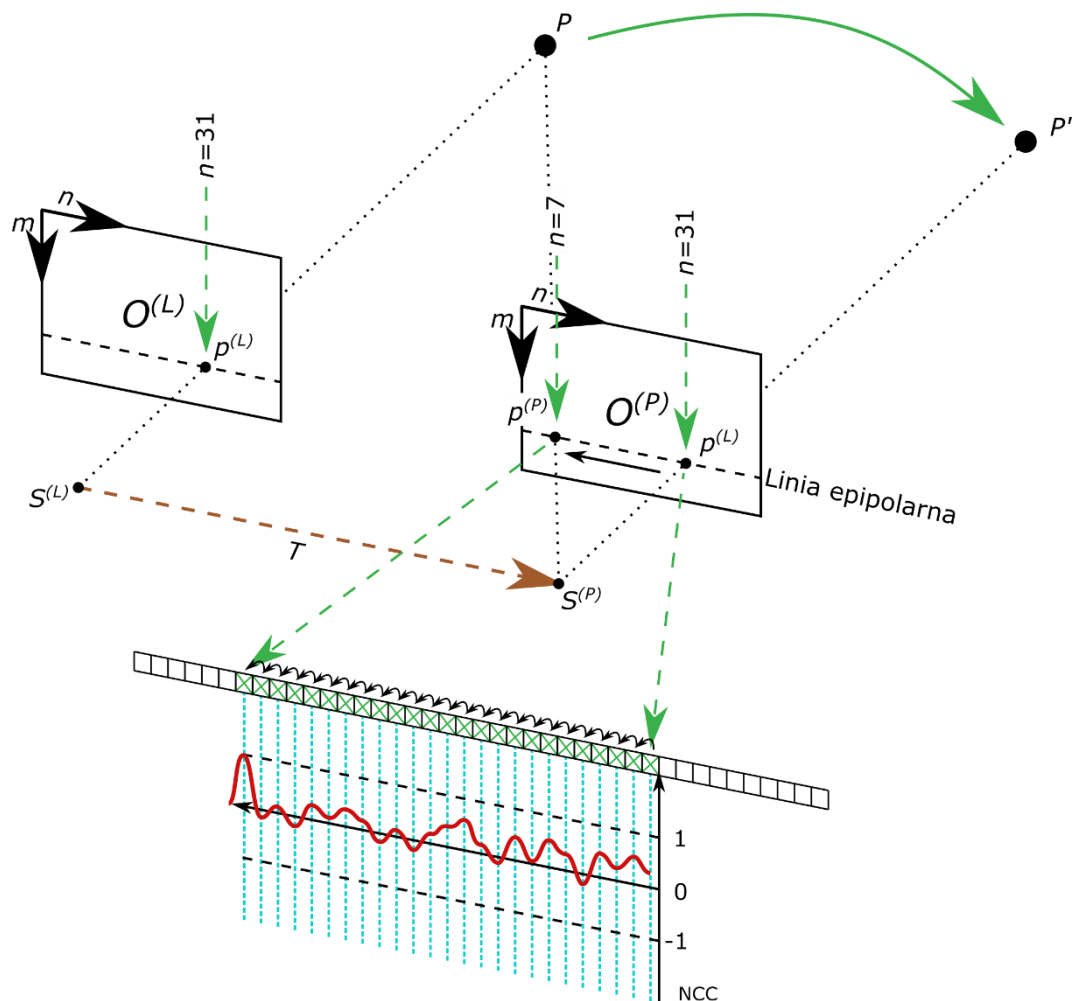
Obliczone w ten sposób dysparycje dla każdego piksela obrazu lewego tworzą macierz dysparycji  $D$ , która razem z elementami orientacji wewnętrznej i zewnętrznej urządzeń obrazujących stanowi argumenty wejściowe do triangulacji.

Osobną grupę metod poszukiwania odpowiedniości obrazów stanowią metody globalne, które są ukierunkowane na jednoczesne wyznaczenie dysparycji dla całego obrazu. Jest to osiągane przez minimalizację funkcji energii technikami optymalizacyjnymi. Zazwyczaj funkcja energii składa się ze składnika odpowiedzialnego za podobieństwo obrazów oraz



ciągłość macierzy dysparycji. Choć złożoność obliczeniowa metod globalnych jest większa, to mają one następujące zalety:

- minimalizacja problemu przesłaniania obiektów,
- efektywniejsze parowanie obszarów homogenicznych,
- wpływ na ciągłość macierzy dysparycji,
- minimalizacja wpływu szumów i zakłóceń.



Rys. 4.8 Parowanie obrazów z wykorzystaniem znormalizowanej korelacji krzyżowej (NCC).

Przykładem stosowanych w stereowizji technik optymalizacyjnych może być metoda oparta o programowanie dynamiczne [36], rachunek wariacyjny [37], rozcięcie grafów [38], relaksacje [39], algorytmy genetyczne [40], sieci neuronowe [41, 42], dyfuzję nieliniową [43], propagację przekonań [44] oraz metody probabilistyczne [45].

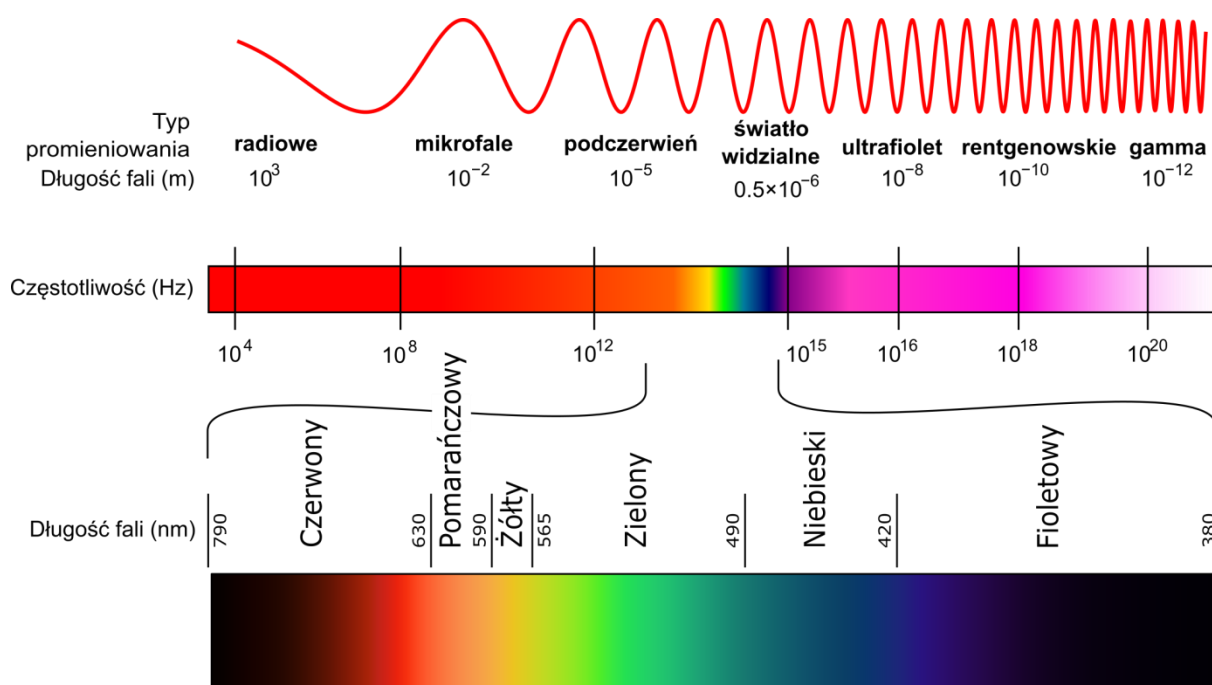
W następnym podrozdziale przybliżony zostanie problem przestrzeni barw, które mogą stanowić bazę dla lokalnych algorytmów blokowych poszukiwania stereokorespondencji. Autor niniejszej rozprawy w dalszej jej części bada również wpływ poszczególnych

składowych przestrzeni barw, jako cech obrazu, na jakość wyników parowania obrazów. Uzasadnione jest więc przybliżenie przestrzeni barw, którą są wykorzystywane w badaniach autora, czyli *RGB*, *HSV* i *Lab*.

## 4.9. Przestrzeń barwne

Podczas tłumaczenia na język polski zagadnień związanych z barwą powstało pewne zamieszanie, związane z odmiennością zwyczajowo używanych określeń [46]. Dlatego na potrzeby niniejszej pracy w tym podrozdziale dokonane zostaną pewne uściślenia powszechnie używanych terminów w odniesieniu do problematyki barwy. Teoria kolorymetrii operuje na wartościach ciągłych, jednak w niniejszej pracy barwa będzie definiowana w sposób dyskretny, zgodnie w powszechnie przyjętym zwyczajem.

Barwa jako najważniejsza cecha obrazu, znalazła się w centrum zainteresowania naukowców już w czasach starożytnych [47]. Problemem barwy zajmował się Pitagoras, czy Arystoteles, a w czasach nowożytnych Isaac Newton, który jako pierwszy podjął próbę usystematyzowania wrażenia barwy, proponując koło barw.



Rys. 4.9 Podział fali elektromagnetycznej na typy promieniowania z wyszczególnieniem pasma widzialnego.

Barwę należy rozumieć dwojako. Jest to odczucie, wrażenie zmysłowe, którego doświadcza się, kiedy fala elektromagnetyczna trafia na siatkówkę, gdzie zostaje przetworzona w impuls elektrochemiczny i nerwem wzrokowym trafia do mózgu. w tym ujęciu nie jest to wielkość fizyczna. Człowiek postrzega jedynie niewielki wycinek spektrum promieniowania elektromagnetycznego nazywany światłem widzialnym. Ewolucja wykształciła w oku fotosensory uczulone na różne długości fal. Widzenie fotopowe, czyli przy dobrych warunkach oświetleniowych, umożliwiają tzw. czopki, zaś widzenie skotopowe, czyli przy minimalnej liczbie fotonów, umożliwiają pręciki. Systematyka fal elektromagnetycznych z wyszczególnieniem pasma widzialnego przedstawiona została na rys. 4.9.

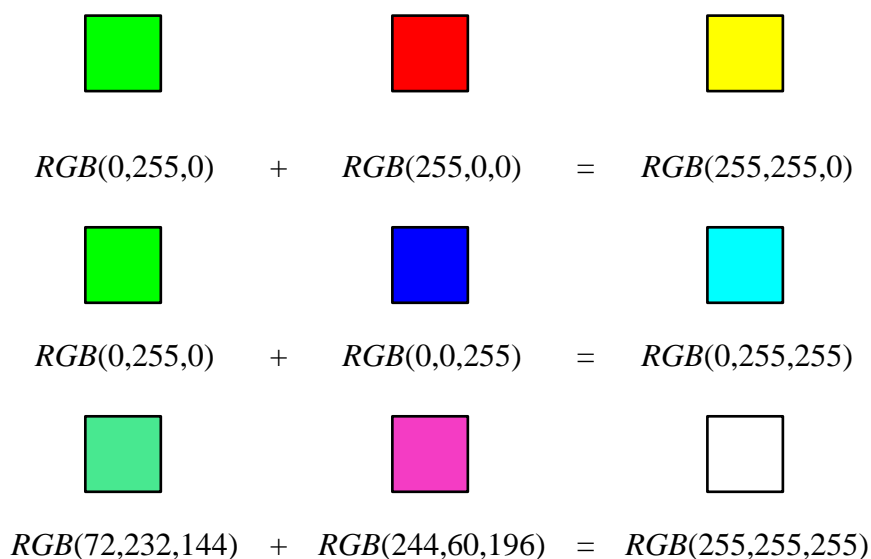
W niniejszej pracy barwa będzie rozumiana jest jako cecha obrazu, wielkość fizyczna w sposób obiektywny wyrażająca odczucie barwy. Kolor (ang. hue) rozumiany jest jako cecha barwy, zwana również odcieniem barwy. Innymi cechami barwy jest również nasycenie (ang. saturation), czy np. jaskrawość (ang. value lub brightness).

W następnym podrozdziale przedstawione zostanie zagadnienie rejestracji barwy zarówno przez ludzkie oko, jak i urządzenia obrazujące.

#### **4.9.1. Widzenie barw**

Rozwój teorii trójkromatyczności dał podstawę do stworzenia modelu przestrzeni barw opartego o trzy barwy podstawowe - czerwoną, zieloną i niebieską. Już wcześniej tworzone były modele bazujące na innych barwach podstawowych, ale dopiero rozwój biochemii i neurofizjologii pozwolił powiązać bezpośrednio model barw ze sposobem odbierania barw przez oko ludzkie. Wcześniej już zostało wspomniane, że widzenie fotopowe jest możliwe dzięki fotoreceptorom zwanym czopkami. w ludzkim oku wyróżniamy trzy rodzaje czopków uczulonych na krótkie długości fal, średnie i długie, w skrócie nazywane SML (ang. *short, medium, long wavelengths*), które, w przybliżeniu, odpowiadają długościom fali odbieranym jako wrażenie barwy niebieskiej, zielonej i czerwonej.

Dzięki trzeciemu prawu Grassmana w teorii barw ma zastosowanie algebra liniowa, co znacznie ułatwiło sformułowanie modelu trójkromatycznego i ułatwia jego stosowanie. Połączenie dwóch barw będzie sumą barw składowych, co przedstawia jest na rys. 4.10.



Rys. 4.10 Przedstawienie trzeciego prawa Grossman'a na przykładzie modelu barw *RGB*.

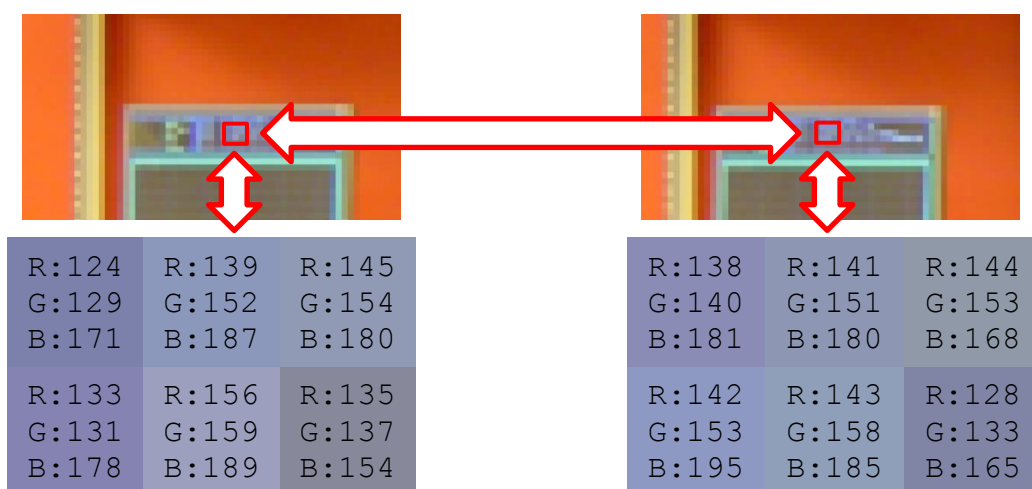
Model *RGB* jest podstawą działania kineskopów, wyświetlaczy LCD, matryc CCD i CMOS. Niestety podobnie jak w przypadku człowieka, zarówno rejestrowanie barwy, jak i jej wyświetlanie, nie jest ustandaryzowane. Każdy człowiek ma uczulone czopki na swoje sobie długości fali, które statystycznie odpowiadają długościom fali odpowiadającym barwą niebieskiej, zielonej i czerwonej, zdefiniowanych jako stałe fizyczne, w granicach pewnego błędu. Różnice w postrzeganiu fali elektromagnetycznej przez poszczególnych ludzi mogą być znaczne.

Problem postrzegania barw jest jednak bardziej skomplikowany, niż to wynika tylko z samych właściwości fotosensorów. Obraz padający na siatkówkę oka, po zamianie na impulsy elektryczne, biegnie przez nerwy wzrokowe do skrzyżowania nerwów [48, 49]. Następnie pasmem wzrokowym dociera do ciała kolankowatego bocznego (LGN) i dalej aksonami w postaci promienistości wzrokowej do pierwotnej kory wzrokowej *Ves1* znajdującej się w obszarze 17 Brodmana'a w płacie potylicznym. Następnie przez obszar *Ves2* dociera do *Ves4* i dalej do struktur asocjacyjnych i motywacyjnych w okolicach kory skroniowej dolnej. Dalej w obszarach skojarzeniowych i wielomodalnych następuje reinterpretacja impulsów reprezentujących barwy. Uszkodzenie któregośkolwiek elementu szlaku wzrokowego lub obszarów funkcji wyższych może prowadzić do ślepoty, w zależności od stopnia uszkodzeń, pomimo prawidłowego funkcjonowania siatkówki. Jednocześnie ustalono, że tylko od 10 do 15 % pobudzeń LGN pochodzi impulsami z siatkówki, zaś całe pozostałe 85 do 90% impulsami z kory wzrokowej!

W szczególności uszkodzenie obszaru *Ves4* prowadzi do achromatopsji czyli ślepoty barw. o innym wrażeniu barwy może również zdecydować zmęczenie, czy nastrój. Duży wpływ na postrzeganie barw ma wiek i stan zdrowia. Również farmakologicznie można zmienić sposób odbioru barw.

Podobnej natury problemy co człowieka, dotyczą urządzeń działających w oparciu o model *RGB*. Ta sama barwa *RGB* (0,0,1) będzie inaczej wyświetlana na kineskopie i inaczej na wyświetlaczu ciekłokrystalicznym, co wynika z różnych charakterystyk zarówno luminoforu, jak i nemetyków. Dodatkowo różnice występują również pomiędzy tymi samymi urządzeniami. Dlatego model *RGB* nazywany jest modelem zależnym od urządzenia (ang. *device dependent model*).

Mimo, że w przestrzeni barw *RGB* powszechnie wykonuje się procedurę poszukiwania odpowiedniości, co przedstawiono na rys. 4.11, to jednak nie nadaje się ona najlepiej do tego celu. Definicja barwy zawiera w sobie jednocześnie kolor, nasycenie i jasność. Chcąc rozjaśnić obraz musimy modyfikować całą barwę, nie znając zmiany saturacji czy jasności. Stanowi to niedogodność, którą starano się wyeliminować opracowując nowe przestrzenie barw.



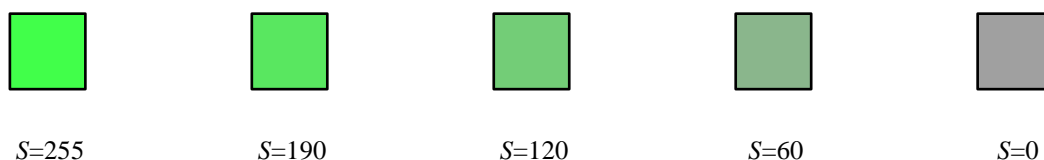
Rys. 4.11 Parowanie obrazów w modelu *RGB*.

O ile model *RGB* instrumentalnie został stworzony na podobieństwo systemu rejestracji barw przez oko ludzkie, to nie jest on zgodny z ludzką percepcją barw. Próbą stworzenia systemu zgodnego z ludzką percepcją barw jest system *HSV* przybliżony w następnym podrozdziale.

### 4.9.2. Percepcja barw

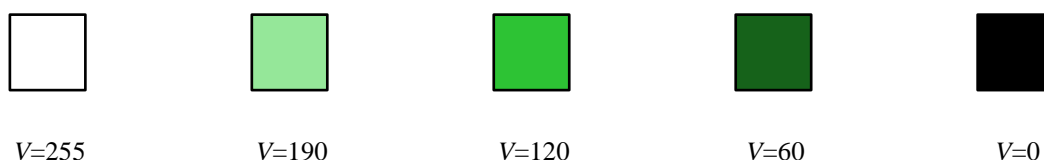
O ile model *RGB* jest ściśle związany z wyświetlaniem barw na konkretnych urządzeniach, czy też ich rejestracją przez konkretne urządzenia, o tyle model *HSV* nawiązuje do sposobu percepcji barw przez człowieka. Źródłem barwy jest światło, najlepiej słoneczne. Barwa istnieje tylko dzięki światłu odbitemu od powierzchni otaczającej nas scenerii. Każda powierzchnia pochłania charakterystyczną dla siebie długość fali, a to co zostanie odbite dociera do oka wywołując wrażenie barwy.

Model *HSV* tworzą trzy składowe – odcień barwy lub kolor (ang. *hue*), nasycenie (ang. *saturation*) oraz jaskrawość (ang. *value* lub *brightness*). Kolor jest kątem liczonym od zera umiejscowionego według konwencji w punkcie czystej czerwieni, liczony przeciwnie do kierunku ruchu wskazówek zegara w mierze stopniowej.



Rys. 4.12 Zmiana barwy w zależności od zmiany nasycenia przy stałym kolorze  $H=87$  i jaskrawości  $V=160$ .

Trzecim składnikiem jest jaskrawość, która określa zmianę jasności barwy, przy niezmienionej wartości nasycenia, co przedstawione zostało na rys. 4.13. Jaskrawość jest odpowiednikiem wielkości fizycznej zwanej luminancją.

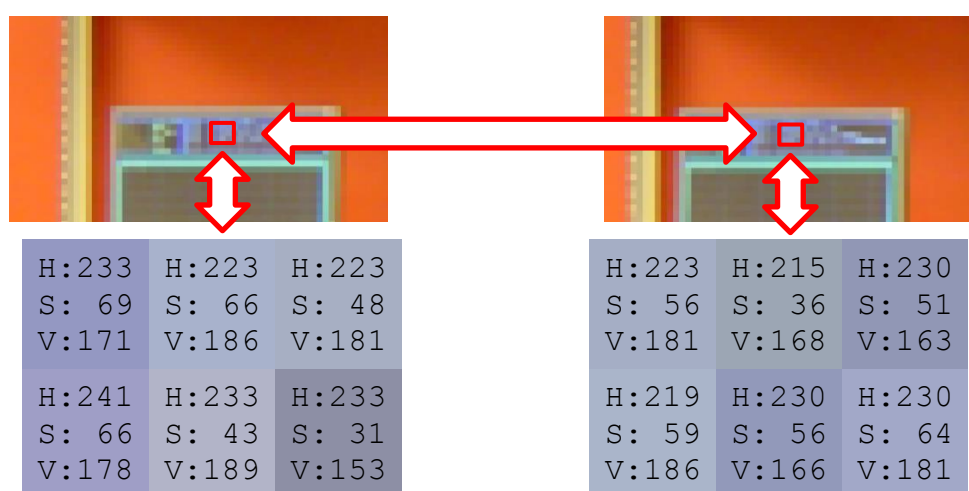


Rys. 4.13 Zmiana barwy w zależności od zmiany jaskrawości przy stałym kolorze  $H=87$  i nasyceniu  $S=160$ .

Czystej czerwieni odpowiada więc wartość 0, czystej zieleni wartość 120, a czystemu niebieskiemu wartość 240. Jeśli zinterpretujemy model *HSV* jako walec, to kolor stanowi obwód podstawy. Promień podstawy stanowi nasycenie wyrażane w przyjętej rozdzielczości, np. od 0 do 255. Nasycenie definiowane jest jako miara udziału czynnika chromatycznego w barwie. Odwołując się do intuicyjnego modelu *RGB*, barwę (238, 156, 118) możemy przedstawić jako sumę dwóch barw (120, 38, 0) i (118, 118, 118). Pierwsza barwa jest składnikiem chromatycznym, czyli pozbawionym udziału światła białego, które stanowi drugą barwę i nosi miano składnika achromatycznego. Dla wartości nasycenia równej zero  $S=0$

pozostaje składnik achromatyczny barwy, czyli odcień szarości, co dobrze oddaje ostatnia pozycja na rys. 4.12.

Model *HSV* umożliwia intuicyjną obróbkę obrazów. Dla przykładu jeśli chcemy rozjaśnić obraz, wystarczy zwiększyć jego wartość *V*, bez ingerencji w wartość koloru *H*, jak to ma miejsce dla modelu *RGB*. Ta sama operacja w modelu *RGB* spowoduje również ujawnienie lub wręcz uwydatnienie szumów zarejestrowanych dla poszczególnych barw podstawowych w stopniu utrudniającym analizę. Dzięki temu model *HSV* znacznie bardziej jest pretendowany do procedury poszukiwania odpowiedniości. Na rys. 4.14 przedstawiono parowanie prostokątnego obszaru w modelu *HSV*. Jest to ten sam obszar parowania, co dla modelu *RGB* przedstawiony na rys. 4.11.



Rys. 4.14 Parowanie obrazów w przestrzeni HSV.

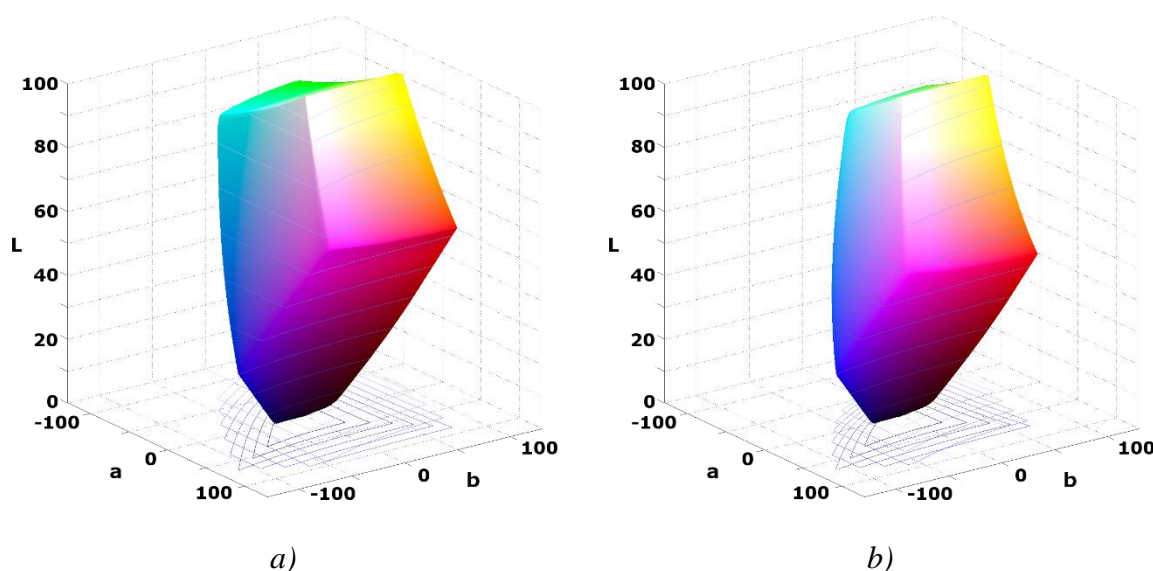
Model *HSV* jest jednak, w mniejszym stopniu niż model *RGB*, wciąż zależny od urządzenia wyświetlającego. Przykładem modelu niezależnego od urządzenia wyświetlającego jest model *Lab*, przedstawiony w następnym podrozdziale.

#### 4.9.3. Model barw niezależny od urządzenia wyświetlającego

Jeszcze na początku XX wieku rozpoczęto prace nad stworzeniem modelu barw niezależnego od urządzeń. w roku 1931 *Commission Internationale de l'Eclairage* ogłosiła model CIE XYZ 1931, który jest podstawą modelu CIE 1976  $L^*a^*b^*$  przedstawionego i wykorzystywanego w niniejszej rozprawie doktorskiej. Nazwa modelu zawiera gwiazdki po każdej literze dla odróżnienia od modelu *Lab* ogłoszonego w 1948 roku przez Hunter'a. Dla

uproszczenia zapisu, w dalszej części niniejszej pracy, ilekroć wystąpi zapis *Lab*, należy przez to rozumieć model CIE 1976  $L^*a^*b^*$ .

Międzynarodowa Komisja Oświetleniowa opracowała model standardowego obserwatora. Stworzone funkcje dopasowujące (ang. *color matching functions*) określały proporcje zmieszania trzech kolorów podstawowych w celu uzyskania, z trzech hipotetycznych receptorów X,Y,Z, odpowiedzi potwierdzającej jednolitość z wyświetlanym światłem monochromatycznym. w ten sposób przestrzeń barw XYZ określono jako trójwymiarową przestrzeń Euklidesową, gdzie każdej barwie odpowiadał wektor  $[x, y, z]^T$ . Opracowany model CIE XYZ pozwala na adresowanie barw z poza zakresu widzialnego dla człowieka.



Rys. 4.15 Reprezentacja gamutów w przestrzeni Lab: a) dla AdobeRGB, przestrzeni barw stworzonej głównie do celów reprodukcji barwnej; b) dla sRGB, przestrzeni barw wspieranej przez urządzenia wyświetlające np. wyświetlacze LCD, czy projektory.

Jednak opracowany model wciąż nie odpowiadał założeniu w myśl którego odległości pomiędzy barwami odpowiadają subiektywnemu odczuciu różnic, jakie odnosi człowiek oglądając porównywane barwy. Obserwator, dwa punkty leżące w niewielkiej odległości od siebie w obszarach barwy niebieskiej lub czerwonej, będzie odbierał jako różne kolory. Ten sam obserwator, w obszarze barwy zielonej nie dostrzeże jednak różnicy pomiędzy dwoma punktami leżącymi w takiej samej odległości.

Różne próby doprowadziły do ogłoszenia 1976 roku modelu CIE Lab, która transformuje model CIE XYZ na trzy współrzędne  $L$ ,  $a$  i  $b$ .



Adobe RGB 1998												
Minimalna wartość							Maksymalna wartość					
	<i>a</i>			<i>b</i>			<i>a</i>			<i>b</i>		
<i>Lab</i>	83.3	<b>-138.0</b>	90.8	33.0	80.3	<b>-109.4</b>	61.4	<b>89.6</b>	75.1	83.3	-138.0	<b>90.8</b>
AdobeRGB	0	255	0	0	0	255	255	0	0	0	255	0
Standard RGB												
Minimalna wartość							Maksymalna wartość					
	<i>a</i>			<i>b</i>			<i>a</i>			<i>b</i>		
<i>Lab</i>	87.7	<b>-86.2</b>	83.2	32.3	79.2	<b>-107.8</b>	53.2	<b>80.1</b>	67.2	87.7	-86.2	<b>83.2</b>
sRGB	0	255	0	0	0	255	255	0	0	0	255	0

Rys. 4.16 Minimalne i maksymalne wartości składowych  $a$  i  $b$  dla gamutów Adobe RGB i sRGB.

Jeśli barwa jest ciemna, to współrzędna  $L$  jest związana z luminancją zależnością liniową. Dla jasnych barw  $L$  jest związane z luminancją zależnością potęgową. Takie zdefiniowanie jaskrawości  $L$  ma pomóc w opisie barw ciemnych, które są słabiej rozróżnialne przez ludzkie oko. Współrzędna chromatyczności  $a$  przedstawia rozkład barwy fuksją (ang. *magenta*) ( $a > 0$ ) lub zielonej ( $a < 0$ ), zaś współrzędna  $b$  rozkład barwy żółtej ( $b > 0$ ) lub niebieskiej ( $b < 0$ ). o ile wartości, jakie może przyjmować jaskrawość  $L$  są ściśle zdefiniowane od 0 do 100, to wartości  $a$  i  $b$  zależą od konkretnego gamutu, czyli zakresu barw które konkretne urządzenie jest zdolne reprodukowac bądź zarejestrować. Dla przykładu dla przestrzeni barw Adobe RGB 1998 minimalne wartości wynoszą  $a = -138$  i  $b = -110$ , a maksymalne  $a = 89,6$  i  $b = 90,8$ . Na rys. 4.15 oraz rys. 4.16 porównano dwa powszechnie używane gamuty AdobeRGB i sRGB.

Barwa obok jasności stanowi domenę poszukiwania odpowiedniości z zastosowaniem funkcji podobieństwa. Poza tymi cechami parować można również inne, takie jak punkty charakterystyczne, krawędzie, czy odcinki w przestrzeni Hough'a, określane jako prymitywy. w następnym podrozdziale przedstawione zostanie zagadnienie punktów charakterystycznych.

#### 4.10. Punkty charakterystyczne

Punkty charakterystyczne na obrazach są często nazywane również narożnikami z powodu powszechnie używanego angielskiego określenia *corners*. Wzięło się ono z faktu, iż

w rzeczywistości fizyczne narożniki znajdujące się na obrazach, dają najsilniejszą odpowiedź detektora w wyniku obliczenia pochodnej po wierszach i kolumnach obrazu.

Najczęściej używanym detektorem punktów charakterystycznych jest detektor Harris'a, który bazuje na lokalnej autokorelacji funkcji sygnału [32], która w rzeczywistości jest sumą kwadratów różnic (ang. *sum squared distance* - SSD) podobnie jak w algorytmie Moravec'a. Algorytm Moravec'a liczy SSD jako podobieństwo pomiędzy oknem o rozmiarach  $(u, v)$  i jego przesunięciem o  $(n, m)$  co jest wyrażone wzorem

$$K_M(n, m) = \sum_u \sum_v w(u, v) \cdot (O(u + n, v + m) - O(u, v))^2. \quad (4.14)$$

W równaniu ważona suma kwadratów różnic pomiędzy dwoma wycinkami obrazu jest oznaczona przez  $K_M(n, m)$ . Mniejsza wartość oznacza większe podobieństwo, a lokalne maksimum wartości oznacza wystąpienie punktu charakterystycznego. Przewagą detektora Harris'a nad Moravec'a jest zastąpienie przesuwanych wycinków obrazu bezpośrednio przez różniczkę detektora po kierunku  $n$  i  $m$ , otrzymując następujące przybliżenie

$$K_H(n, m) \approx \sum_u \sum_v w(u, v) \cdot (O_n(u, v)n - O_m(u, v)m)^2. \quad (4.15)$$

W powyższym równaniu częściowe pochodne funkcji obrazu są oznaczone jako  $O_n$  i  $O_m$ . Ta modyfikacja wpłynęła na zredukowanie złożoności obliczeniowej algorytmu. Jeśli zastosujemy dodatkowo maskę kołową lub maskę kołową ważoną (funkcja Gaussa) odpowiedź będzie izotropowa. Ponieważ formuła (4.15) wciąż zbyt łatwo odpowiadała na krawędzie, przeformułowano algorytm wprowadzając symetryczną macierz  $A$ , nazwaną macierzą Harris'a i zdefiniowano operator w oparciu o wartości własne  $\lambda_1$  i  $\lambda_2$  tej macierzy w następujący sposób:

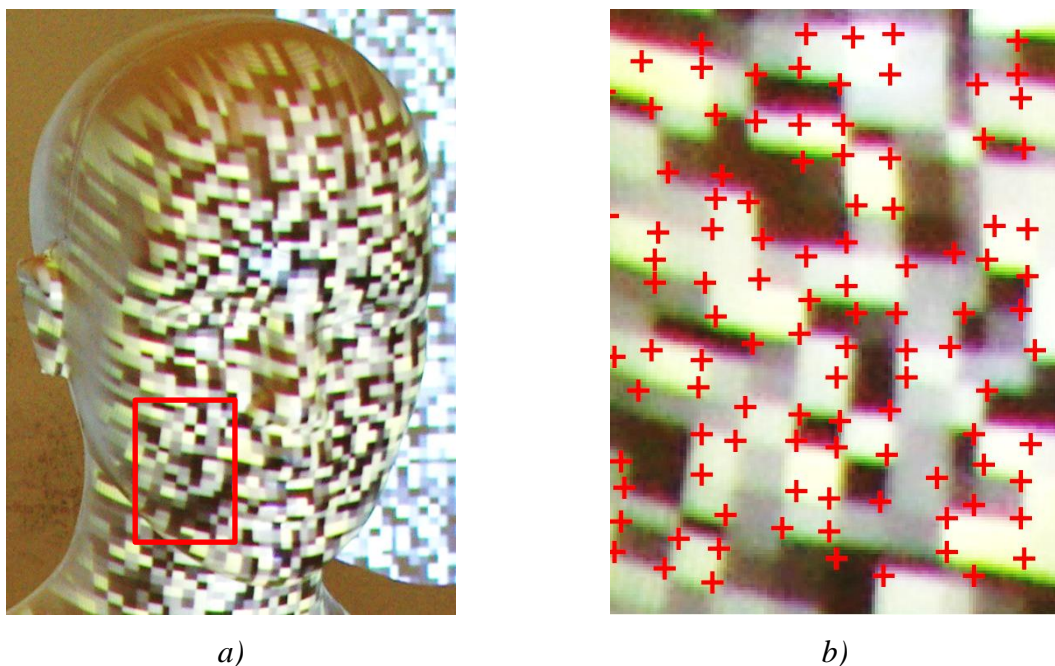
$$K_H = \lambda_1 \cdot \lambda_2 - \kappa \cdot (\lambda_1 + \lambda_2)^2 = \det(A) - \kappa \cdot \text{trace}^2 A. \quad (4.16)$$

Ostatecznie wartości własne macierzy Harris'a zostały zastąpione wyznacznikiem macierzy a i śladem macierzy  $A$ , co zmniejszyło złożoność czasową algorytmu. Przykład detekcji punktów charakterystycznych metodą Harris'a przedstawiono na rys. 4.17.

By wyeliminować problematyczny parametr  $\kappa$ , który trzeba każdorazowo eksperymentalnie weryfikować, w pracy [33] zaproponowano nową modyfikację detektora Harris'a, która nie zawiera żadnego parametru, w postaci

$$K_N = \frac{\det(A)}{\text{trace}(A) - \varepsilon}. \quad (4.17)$$

Stała  $\varepsilon$  jest bliska zera, by uniknąć jedności w mianowniku w wypadku zerowego rzędu macierzy Harris'a  $A$ .



Rys. 4.17 Detekcja punktów charakterystycznych z użyciem operatora Harris'a: *a)* zobrazenie twarzy przygotowane do wykrywania narożników; *b)* wykryte punkty charakterystyczne w zaznaczonym na obrazie *a)* prostokącie.

W pracy [50] przeformułowano poszukiwanie punktów charakterystycznych w zagadnienie śledzenia punktów pomiędzy następującymi po sobie ramkami obrazu. Matematyczne podstawy są bardzo podobne do omawianego wcześniej detektora Harris'a, z tą różnicą że macierz Harris'a jest teraz macierzą autokorelacji. Funkcja nazwana od naukowców, którzy ją opublikowali, funkcją Shi-Tomasiego, jest najmniejszą wartością własną macierzy autokorelacji a wyrażona równaniem:

$$K_{ST} = \lambda_{\min} \cdot (A). \quad (4.18)$$

Metoda została ponownie sformułowana w pracy [51] przez Tomasi i Kanade i dlatego funkcjonują dwie nazwy tej metody Kanade-Tomasi i Shi-Tomasi.

W następnym podrozdziale przedstawione zostanie zagadnienie wykrywania krawędzi, jako następnej cechy, która może stanowić podstawę dla określenia odpowiedniości obrazów.

#### 4.11. Wykrywanie krawędzi

Wykrywanie krawędzi polega na detekcji zmiany luminancji obrazu wzdłuż wydłużonego, względnie liniowego obszaru. Matematycznie zmianę luminancji wykrywamy obliczając pierwszą lub drugą pochodną obrazu. w praktyce aproksymujemy pochodną obrazu wykonując jego konwolucję z odpowiednio przygotowaną maską. Większość praktycznie wykorzystywanych metod, takich jak Canny, Sobel, Prewitt, Roberts czy Scharr bazuje na badaniu pierwszej pochodnej, nazywanej również gradientem luminancji lub wprost gradientem obrazu.

Przy nagłej zmianie luminancji gradient obrazu osiąga ekstremum lokalne, co zostało przedstawione na rys. 4.19. Tempo zmiany gradientu luminancji przedstawia druga pochodna. Lokalne ekstremum pierwszej pochodnej określone jest przez przejście przez zero drugiej pochodnej, co jest również dobrym kryterium podczas poszukiwania krawędzi.

Operator Sobela [52] aproksymuje gradient obrazu dokonując splotu dwóch masek o rozmiarze  $3 \times 3$  elementy z obrazem  $O$ , osobno po kolumnach  $n$  i po wierszach  $m$ . w wyniku otrzymujemy dwa obrazy  $G_n$  i  $G_m$  będące pochodnymi cząstkowymi obrazu  $O$ , co można zapisać jako:

$$G_n = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * O \quad (4.19)$$

$$G_m = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * O \quad (4.20)$$

gdzie  $*$  oznacza operację dwuwymiarowego splotu.

Inną maskę zaproponował Prewitt [53] zdefiniowaną jako:

$$G_n = \begin{bmatrix} -1 & 0 & +1 \\ -1 & 0 & +1 \\ -1 & 0 & +1 \end{bmatrix} * O \quad (4.21)$$

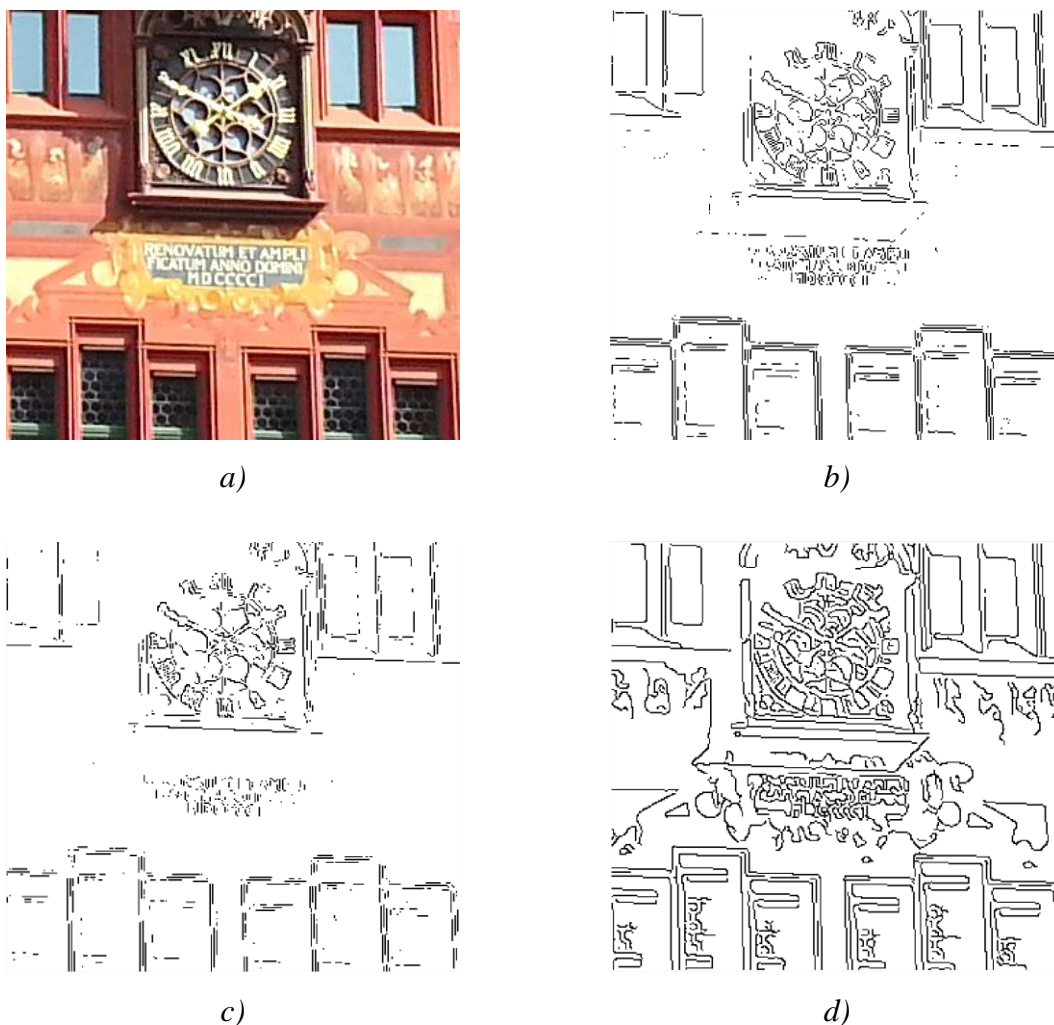
$$G_m = \begin{bmatrix} +1 & +1 & +1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} * O. \quad (4.22)$$

Operator Roberta [52] zaproponowany najwcześniej stosuje maskę 4-elementową postaci:

$$G_n = \begin{bmatrix} +1 & 0 \\ 0 & -1 \end{bmatrix} * O \quad (4.23)$$

$$G_m = \begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix} * O. \quad (4.24)$$

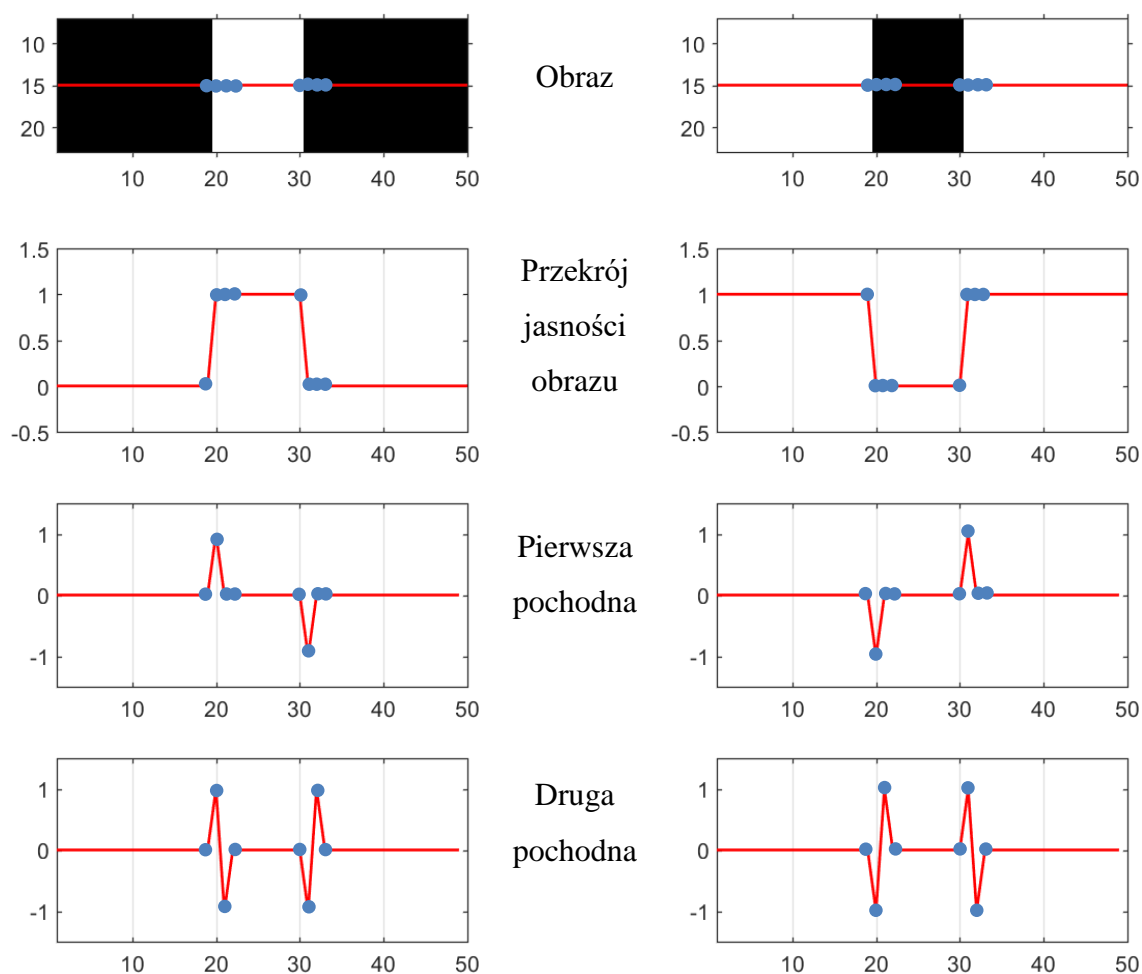
Przykłady działania detektorów krawędzi przedstawione zostały na rys. 4.18.



Rys. 4.18 Wynik działania detektorów krawędzi; a) oryginalny obraz; b) detektor Sobel'a; c) detektor Roberta; d) detektor Canny'ego.

Próg binaryzacji ustalony został dla operatora Sobel'a na 0,2124, dla operatora Roberts'a na 0,2378, natomiast dla Canny próg minimalny na 0,0750 oraz próg maksymalny na 0,1875.

Operator Canny'ego [54] jest złożoną metodą wieloetapową, która daje wciąż najlepsze efekty, dla zastosowań ogólnych. Składa się ona z czterech etapów. Najpierw redukowany jest szum funkcją Gaussa. Następnie aproksymowany jest gradient obrazu. Linie grubsze niż 1 piksel są zwężane, usuwając mniej ważne piksele.



Rys. 4.19 Kształt pierwszej i drugiej pochodnej obrazu w zależności od zmian luminancji obrazu. Na niebiesko oznaczono punkty pomiarowe.

W ostatnim etapie następuje binaryzacja obraz stosując progowanie z histerezą. Metoda Canny'ego wykorzystuje rachunek wariacyjny do znalezienia funkcji, która optymalizuje dany funkcjonal, czyli zbiór założeń. Funkcja ma spełniać trzy warunki. Pierwszym jest efektywność, czyli detektor ma znaleźć wszystkie krawędzie

Drugim warunkiem jest dokładność, czyli znaleziona krawędź ma znajdować się najbliżej rzeczywistej krawędzi. Trzecim warunkiem jest brak redundancji - linia ma być znaleziona

tylko raz, bez względu na poziom szumów w obrazie. Optymalna funkcja może być aproksymowana pierwszą pochodną funkcji Gaussa zgodnie ze wzorem

$$B = \frac{1}{159} \begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix} * O \quad (4.25)$$

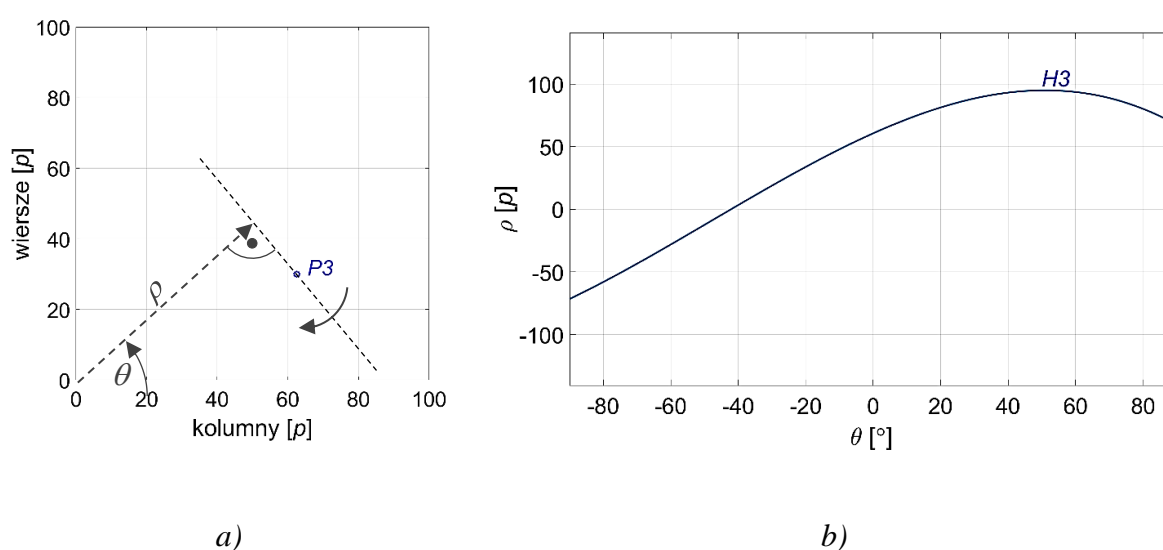
dla  $\sigma = 1,4$ .

Poszukiwanie krawędzi z wykorzystaniem drugiej pochodnej wiąże się obliczeniem Laplasjanu dla obrazu o [53]. Laplasjan jest jednak bardzo czuły na szum, co utrudnia procedurę wykrywania krawędzi. By wyeliminować szum wygładza się obraz stosując filtr rozmycia w postaci funkcji Gaussa.

Następną cechą podlegającą parowaniu, która zostanie przedstawiona w następnym podrozdziale, jest domena odcinków wyliczonych z transformaty Hough'a.

## 4.12. Transformata Hough'a

Opracowana w roku 1952 przez Paula Hough'a [55] transformata pozwala wykryć na obrazach binarnych liniowe skupiska pikseli. Nadaje się więc do pozyskania cech obrazów, które można poddać dalszej analizie.



Rys. 4.20 Transformata Hough'a: a) dla punktu  $P3$  w przestrzeni danych; b) odpowiadająca punktowi krzywa sinusoidalna w przestrzeni parametrów.

W odróżnieniu od detektorów krawędzi, transformata Hough'a (TH) w ujęciu klasycznym pozwala wykrywać jedynie odcinki prostoliniowe. Zasadnicza różnica dotyczy oczywiście działania obu metod. Detektory krawędzi obliczają pierwszą bądź drugą pochodną obrazu jako np. funkcji jasności. Następnie wynik jest binaryzowany z zastosowaniem odpowiedniego progu. Transformata Hough'a każdemu punktowi w przestrzeni obrazu, nazywanej tu przestrzenią danych, przyporządkowuje krzywą sinusoidalną w przestrzeni parametrów. Jako punkt należy tu oczywiście rozumieć bit obrazu o stanie 1. Jeśli w przestrzeni danych poprowadzimy przez punkt  $P3$  przedstawiony na rys. 4.20 prostą o równaniu

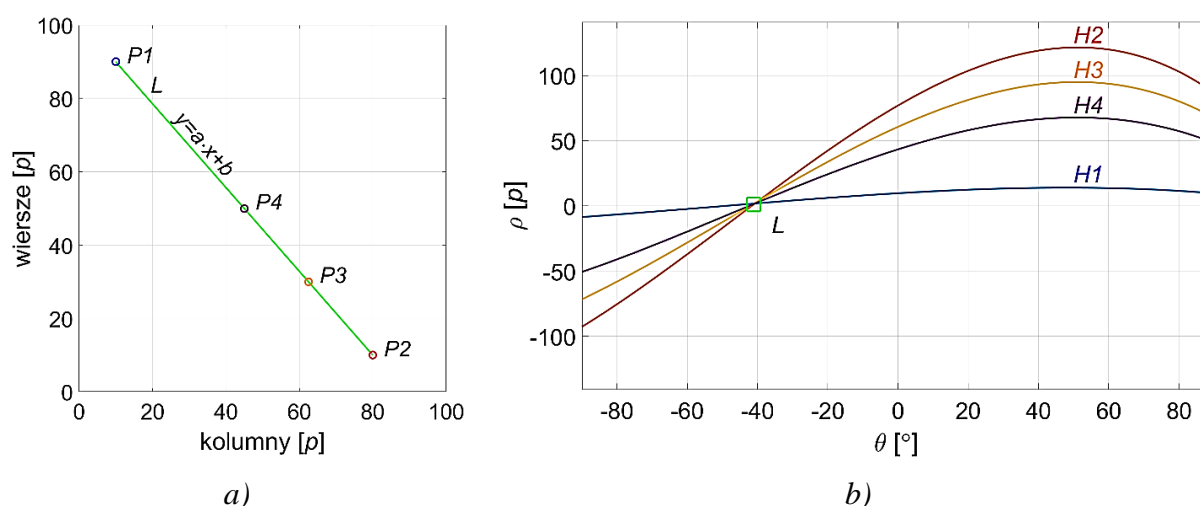
$$y = a \cdot x + b \quad (4.26)$$

i obrócimy ją wokół tego punktu, to otrzymane wartości  $\rho$  i  $\theta$  w przestrzeni parametrów utworzą krzywą  $H3$  o równaniu

$$\rho = x \cdot \cos\theta + y \cdot \sin\theta \quad (4.27)$$

gdzie  $\rho$  oznacza odległość prostej od początku układu współrzędnych w przestrzeni danych, wyznaczanej wzdłuż normalnej do prostej, a  $\theta$  oznacza kąt pomiędzy normalną do prostej, a osią  $x$ .

Jeśli kilka punktów jest ze sobą odpowiednio skorelowane, jak to widać na rys. 4.21, to odpowiadające im krzywe w przestrzeni parametrów przetną się w jednym punkcie, którego współrzędne  $\theta$ ,  $\rho$  określają kąt do normalnej i odległość do prostej  $L$  w przestrzeni danych  $x$ ,  $y$ .



Rys. 4.21 Transformata Hough'a: a) dla odcinka  $P1P2$  zawierającego punkty  $P3$  i  $P4$ , jest podzbiorem prostej o równaniu  $y=ax+b$ ; b) odpowiadający punktowi pęk krzywych sinusoidalnych w przestrzeni parametrów przecinających się w jednym punkcie, odpowiadającym kątowi  $\theta$  i odległości normalnej  $\rho$  do prostej  $L$  w przestrzeni danych.

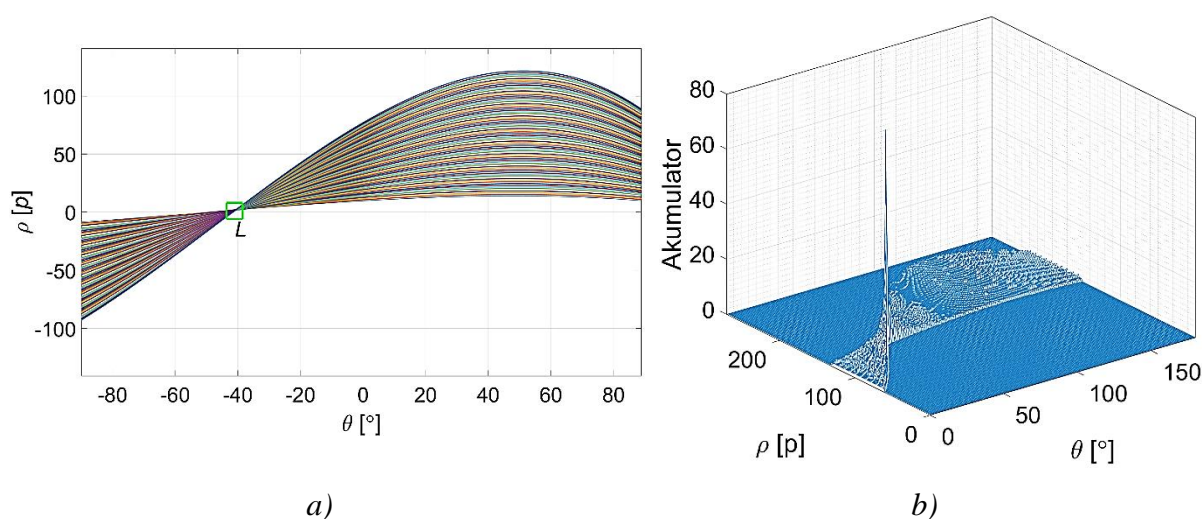
Jeśli punkty są ze sobą słabo skorelowane, to punkt przecięcia się odpowiadających im krzywych sinusoidalnych jest rozmyty. Krzywe przecinają się wtedy w wielu punktach



umiejscowionych bardzo blisko siebie, przy czym punkty przecięcia w przestrzeni parametrów odpowiadają grupom punktów lepiej ze sobą skorelowanych w przestrzeni danych. Im bardziej punkt przecięcia jest rozmyty tym z mniejszą dokładnością określone są parametry  $\theta$  i  $\rho$ . Jak widać na rys. 4.20 w klasycznej notacji kąt  $\theta$  mierzy się od osi  $x$ . We współczesnych implementacjach TH [56] zapożyczono sposób definiowania układu współrzędnych od systemu stosowanego dla obrazów bitmapowych. Początek układu współrzędnych przyjmuje się w górnym lewym narożniku obrazu. Oś  $x$  pokrywa się z górną krawędzią obrazu, a oś  $y$  z lewą krawędzią. Wartości  $\theta$  i  $\rho$  znajdujące się pod osią  $x$  przyjęto oznaczać jako wartości dodatnie, zaś te leżące ponad osią  $x$  jako wartości ujemne.

Dla większej liczby punktów na obrazie otrzymujemy wykres w przestrzeni parametrów w formie smugi sinusoidalnej, przedstawionej na rys. 4.22, która składa się z krzywych sinusoidalnych.

Jeśli na obrazie występuje liczba prostych  $p$ , to w przestrzeni parametrów otrzymamy splecione ze sobą wstęgi sinusoidalne w liczbie większej od  $p$ , których krzywe przecinają się w dokładnie  $p$  punktach. Zjawisko wzajemnego splotu wiązek krzywych sinusoidalnych wynika z faktu, iż przez każde dwa punkty na obrazie można przeprowadzić prostą, choć na obrazie nie są one częścią wspólnej krawędzi.



Rys. 4.22 Transformata Hough'a dla zbioru punktów tworzących prostą  $L$ : a) krzywe sinusoidalne tworzące wstęgę przecinają się dokładnie w jednym punkcie, którego współrzędne  $\theta$  i  $\rho$  opisują prostą w przestrzeni danych; b) przestrzenna wizualizacja wartości akumulatora, czyli liczby krzywych sinusoidalnych przechodzących przez oczko siatki akumulacyjnej.

Dodatkowo na takiej prostej mogą się znaleźć inne przypadkowe punkty. Na rys. 4.22 widać, oprócz silnego lokalnego maksimum, iż wstęga nie jest płaska, lecz jest w każdym miejscu wypukła. Oś  $z$  reprezentuje akumulator, czyli sumę krzywych sinusoidalnych przechodzących

przez oczko przestrzeni parametrów o ustalonym rozmiarze. Dla realistycznych obrazów, np. fotografii, wartość akumulatora będzie zawsze większa od wartości 2. w celu określenia parametrów  $\rho$  i  $\theta$  można posłużyć się histogramem tablicowym. Przestrzenna wizualizacja takiego histogramu przedstawiona jest na rys. 4.22. Przestrzeń parametrów dzieli się na siatkę o określonych wymiarach oczka  $\Delta\rho$  i  $\Delta\theta$ . Oczka siatki nazywamy akumulatorami, ponieważ sumuje się w nich liczbę przechodzących przez nią krzywych sinusoidalnych. Rozmiar akumulatora musi być zoptymalizowany i przyjęło się stosować wartości  $\Delta\rho = 2,5$  piksela i  $\Delta\theta = 2,5^\circ$ . Dla zbyt małego rozmiaru oczka siatki lokalne maksima ulegną rozmyciu, zaś dla zbyt dużego rozmiaru mogą wystąpić fałszywe maksima. w oby wypadkach trudne będzie precyzyjne ustalenie parametrów  $\rho$  i  $\theta$ .

Po ekstrakcji cech, można przystąpić do ich korelowania pomiędzy obrazami stereopary. Jedną z najbardziej elementarnych metod poszukiwania stereokorespondencji, a zarazem bardzo skuteczna opiera się na funkcji podobieństwa powszechnie znanej jako autokorelacja. Zostanie ona przybliżona w następnym podrozdziale.

#### 4.13. Znormalizowana korelacja krzyżowa

Klasyczną funkcją podobieństwa stosowaną w stereowizji jest znormalizowana korelacja krzyżowa *NCC* [34, 57] (ang. *normalized cross corelation*) czasami nazywana autokorelacją dana wzorem:

$$O_{corr} = \frac{\sum_m^M \sum_n^N (O(m, n) - \bar{O}) \cdot (H(m, n) - \bar{H})}{\sqrt{\sum_m^M \sum_n^N (O(m, n) - \bar{O})^2 \cdot \sum_m^M \sum_n^N (H(m, n) - \bar{H})^2}} \quad (4.28)$$

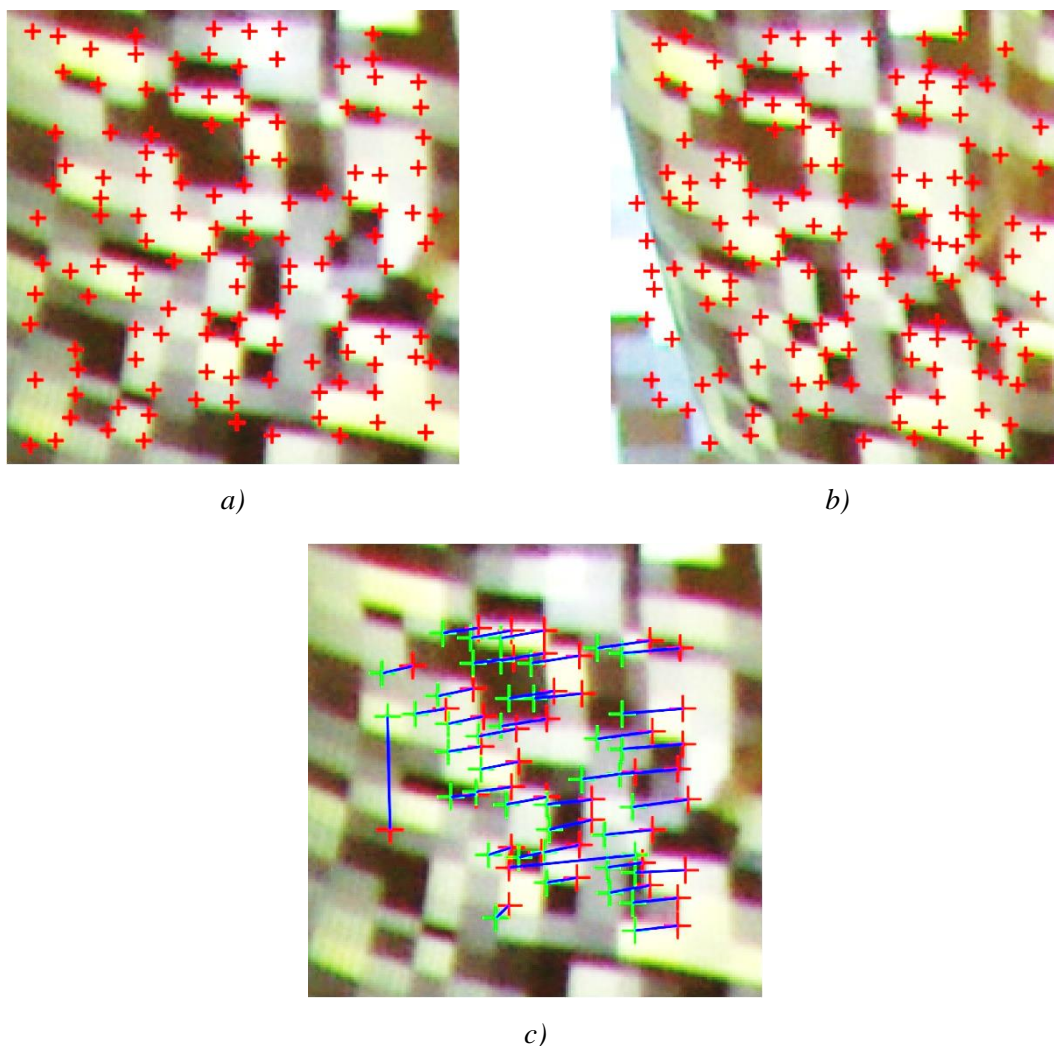
gdzie:

- $O(m, n)$  - jasność piksela  $m, n$  obrazu  $O$ ,
- $H(m, n)$  - jasność piksela  $m, n$  maski  $H$ ,
- $\bar{O}$  - średnia jasność obrazu,
- $\bar{H}$  - średnia jasność maski.

Dzięki subtrakcji wartości średniej jasności obrazu i maski otrzymujemy znormalizowany obraz wzorca i przeszukiwanego tła. Mianownik w postaci wariancji nie tylko daje wygodny do analizy wynik w zakresie od -1 do +1, ale i pozwala, w niewielkim zakresie,

na pracę ze zrotowanymi względem obrazu maskami. Wynikiem oczekiwanym jest  $+1$ , natomiast  $-1$  jest wynikiem skorelowania 100 % nasycenia z 0 % nasyceniem, czyli pozytywu z negatywem.

Wynik poszukiwania odpowiedniości metodą autokorelacji przedstawiony został na rys. 4.23. Maskę ustawiono została na  $51 \times 51$  pikseli. Na rys. 4.23 *a)* przedstawiono czerwonymi markerami wykryte punkty charakterystyczne na lewym obrazie, a na *b)* na prawym obrazie. Na *c)* czerwonymi markerami oznaczono punkty charakterystyczne wykryte na obrazie lewym, natomiast zielonymi markerami na obrazie prawym. Niebieską linią połączone zostały odpowiadające sobie punkty. Linie te stanowią jednocześnie promienie rdzenne.



Rys. 4.23 Stereokorespondencja obrazów; *a)* punkty charakterystyczne na lewym obrazie; *b)* punkty charakterystyczne na prawym obrazie; *c)* punkty odpowiadające sobie połączone są niebieską linią, czerwone krzyże reprezentują punkty homologiczne na lewym obrazie, a zielone punkty na prawym.

Zastosowana metoda eliminacji błędnych sparowań, nie gwarantuje 100 % poprawności wyniku. Wyraźnie widać jedną linię pionową, która reprezentuje błędne sparowanie, pomimo tego, że błędne sparowania zostały wcześniej odfiltrowane korzystając z metody RANSAC [58]. Autokorelacja jest wykorzystywana w metodach lokalnych. Przykładem metody globalnej jest metoda oparta o rachunek wariacyjny, przedstawiona w następnym podrozdziale.

#### 4.14. Metoda wariacyjna

Można przyjąć, iż wartości dysparycji dla poszczególnych pikseli tworzą pole paralaksy, które jest przykładem pola wektorowego. Takie pole wektorowe często jest nazywane w analizie sekwencji video polem przepływu optycznego (ang. *optic flow field*). Prace [59] oraz [60] dowodzą, iż metoda wariacyjna (ang. *variational* - VAR) należy do najbardziej efektywnych metod rekonstrukcji przepływu optycznego. Polega ona na minimalizacji funkcjonału energii po przez rozwiązanie odpowiedniego równania Eulera-Lagrange'a, które jest reprezentowane przez system równań cząstkowych różniczkowych [37, 61].

Przyjmijmy, że skalar natężenia światła jest zachowany w macierzy pikseli i  $O(m,n)$  to współrzędne piksela w prostokątnej domenie obrazu. Mając dwa obrazy tworzące stereogram oznaczmy pierwszy obraz przez  $O^{(L)}$  i drugi przez  $O^{(P)}$ . Następnie dokonujemy estymacji przesunięcia piksela z pierwszego obrazu do nowej pozycji na drugim obrazie. Taka estymacja ma sens zakładając, że te same szczegóły znajdują się na obu obrazach. Tymi szczegółami mogą być wartość jasności, wyższe pochodne obrazu (takie jak gradient albo Hessian) lub wyrażenia skalarne (norma gradientu, Laplacian, wyznacznik Hessianu) [62]. Dla ułatwienia dalszego wywodu przyjmijmy stałość wartości jasności.

Jeśli  $(m,n)$  jest współrzędną piksela na pierwszym obrazie i  $u(m,n)$  jest miarą dysparycji, to  $(m, n+u(m,n))$  jest nową pozycją piksela na drugim obrazie. Formułując problem w ten sposób możemy stwierdzić, że wyznaczenie pola paralaksy jest właściwie niczym innym, jak wyznaczeniem wektora pola  $(0, u(m,n))$ . Teraz zapiszmy założenie dotyczące wartości jasności pikseli na obu obrazach:

$$O^{(L)}(m,n) - O^{(P)}(m, n + u(m,n)) = 0 \quad (4.29)$$

Ponieważ obiekty podlegające rekonstrukcji 3D mają charakter przestrzenny, często z silnie zróżnicowanymi paralaksami, określone punkty widoczne na obrazie pierwszym

niekoniecznie będą widoczne na obrazie drugim, gdyż naturalnie będą zasłonięte przez kształt obiektu. Problematiczne są również nieprzewidywalne właściwości odbijania światła przez różne powierzchnie, powodujące silne refleksy. z tego powodu trudno będzie dotrzymać równość (4.29). Niemniej jednak możemy równość (4.29) łatwo zmodyfikować do formy

$$O^{(L)}(m, n) - O^{(P)}(m, n + u(m, n)) \rightarrow 0 \quad (4.30)$$

i wtedy funkcjonal energii, bazując na warunku zachowania stałości jasności punktu, może być zapisany jako

$$Z(u(m, n)) = \sum_{m=1}^M \sum_{n=1}^N \|O^{(L)}(m, n) - O^{(P)}(m, n + u(m, n))\|^2 \rightarrow 0. \quad (4.31)$$

Celem metody wariacyjnej jest znalezienie funkcji  $u(m, n)$  która minimalizuje funkcjonal energii  $Z(u(m, n))$ . Innymi słowy mając dany funkcjonal energii musimy go zminimalizować, by znaleźć najlepsze rozwiązanie wartości pola dysparycji. Jeśli dany funkcjonal jest wypukły będzie na pewno miał unikalne rozwiązanie, które go zminimalizuje.

W następnym podrozdziale przedstawiona zostanie metoda probabilistyczna.

#### 4.15. Metoda bazująca na informacji wzajemnej.

Zakładając, iż obrazy są zrektyfikowane, funkcję kosztu dla odpowiedniości punktów charakterystycznych można definiować jako informację wzajemną  $MI$  [45, 63] (ang. *mutual information* - MI) daną wzorem

$$MI = H_{O^{(L)}} + H_{O^{(P)}} - H_{O^{(L)}O^{(P)}} \quad (4.32)$$

gdzie  $H_O$  – entropia stereogramów,  
 $H_{O^{(L)}O^{(P)}}$  – entropia wzajemna.

Tak zdefiniowana funkcja jest w dużej mierze niezależna od zmian oświetlenia sceny pomiędzy zobrazowaniami.

Zarówno entropię stereogramów, jak i entropię wzajemną definiuje się z rozkładu prawdopodobieństwa jasności dla obrazów.

W przypadku stereowizji entropia wzajemna  $H_{O(L)O(P)}$  może być bardzo mała ponieważ oba obrazy są od siebie silnie zależne, co daje nam dużą wartość informacji wzajemnej  $MI$ . Poszukiwanie właściwej wartości dysparycji, polegać tu będzie na przemieszczaniu obrazu prawego względem lewego i wyliczaniu każdorazowo wartość entropii wzajemnej. Znalezienie najniższej wartości entropii wzajemnej oznacza znalezienie najlepszej wartości dysparycji.

Bazując na powyższym, problem poszukiwania stereokorespondencji można zdefiniować, jako poszukiwanie macierzy dysparycji  $D$ , która minimalizuje energię  $Z(D)$ . Energię definiuje się w oparciu o funkcję kosztu wygładzoną warunkami opartymi o różnice pomiędzy kolejnymi wartościami dysparycji. Niestety globalna minimalizacja funkcji energii  $Z(D)$  jest problemem NP-zupełnym z powodu nieciągłości macierzy dysparycji [64]. Problem można uprościć przeprowadzając minimalizację wzdłuż linii epipolarnych [65]. Jeszcze lepsze wyniki osiąga się przeprowadzając minimalizację z kilku kierunków [63], co najmniej 8, a najlepiej 16.

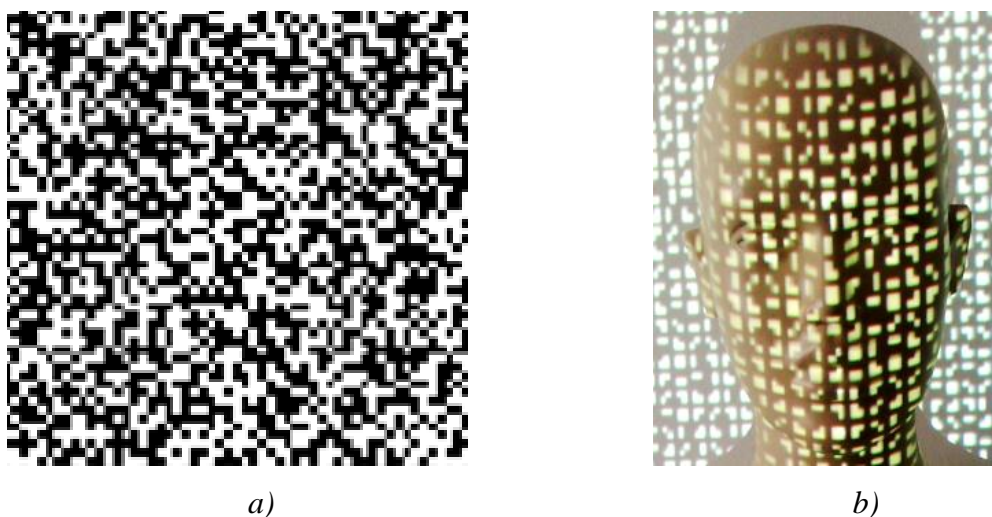
W następnym podrozdziale, bazując na dotychczasowych przeprowadzonych badaniach autora niniejszej rozprawy, przedstawione zostaną niektóre aspekty techniczne towarzyszące poszukiwaniu stereokorespondencji.

#### **4.16. Optymalizacja procesu poszukiwania odpowiedniości**

Sterując parametrami paralaksy, takimi jak wartość minimalna oraz zakres paralaksy, można kontrolować, znaną z optyki, głębię ostrości, tyle że w przestrzeni przedmiotowej. Dla punktów położonych w nieskończoności dysparycja jest równa zero, a w miarę zbliżania się do środka rzutów rośnie. Ustalając minimalną paralaksę można od razu zawęzić głębię ostrości, odrzucając daleki plan opracowywanych zobrażeń, co również znacznie obniża koszt obliczeniowy procedury. Głębię ostrości można oczywiście powiększyć zmniejszając bazę  $T$  pomiędzy zobrazeniami, zmniejszając wielkość ogniskowej  $f$ , zwiększając zakres dysparycji, który będzie przeszukiwany oraz po przez zwiększenie rozdzielczości obrazu, czyli zwiększając liczbę punktów podlegających stereokorespondencji.

Przed przystąpieniem do poszukiwania odpowiedniości, szczególnie metodami lokalnymi, zalecane jest przeprowadzenie wstępnej obróbki obrazu, celem ulepszenia ich jakości, np. po przez zastosowanie funkcji normalizacji. Często skutkuje to poprawieniem jakości stereokorespondencji.

Dla zwiększenia rozdzielczości chmury punktów można zastosować wzbogacenie powierzchni jednorodnych, do których zalicza się ciało ludzkie, po przez projekcję wysoko kontrastowego wzorca na elementy sceny. Jedną z możliwych tekstur przedstawioną na rys. 4.24.a złożoną z 10 elementarnych bloków zbudowanych z elementów  $3 \times 3$  piksele zaproponowano w [66].



Rys. 4.24. Wzorec wzbogacający powierzchnie jednorodne: *a)* macierz złożona z bloków  $3 \times 3$  pola rozmieszczonych zgodnie z quasi-normalnym rozkładem  $\pi$ ; *b)* zniekształcenie rzutowanego wzorca, jako wynik deniwelacji obiektu.

Dzięki quasi-normalnemu rozkładowi liczby  $\pi$  [67–69] zgodnie z którym zostały rozmieszczone elementarne bloki, zachodzi pewność niepowtarzalności układu bloków. Jednocześnie, dysponując nawet małym rozpoznany obszarem wzorca, możemy łatwo odtworzyć część brakującą lub zlokalizować pozycję w całości.

Na rys. 4.24.b przedstawiono deniwelację obiektu, na przykładzie np. okolic nosa lub uszu, która powoduje brak widoczności pewnych obszarów na obu obrazach, co uniemożliwia znalezienie pełnej ich odpowiedniości. Trudność w parowaniu sprawiają również obszary pozostające w cieniu lub w silnym błysku odbitego światła [70]. Zobrazowanie takich elementów wykonane z dwóch różnych geometrycznie punktów sceny, po konwersji na odcienie szarości lub zbinaryzowaniu, są trudne do skorelowania. Pod wpływem czynnika zakłócającego, ten sam element posiada różne odwzorowania na obu obrazach na stereoparze. Często element pozostaje pod wpływem czynnika zakłócającego tylko na jednym obrazie. w wyniku otrzymane korelacje zawsze zawierają fałszywe sparowania, które należy odfiltrować. Jedną z metod, którą można zastosować do pozbycia się błędnych sparowań jest RANSAC [71], której principia zostaną po krótkce przedstawione.

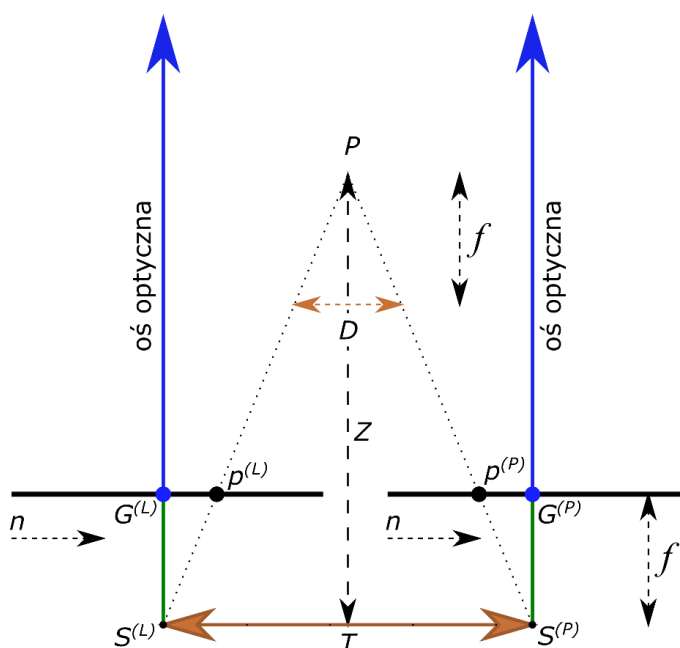


Ze zbioru wstępnie wytypowanych par punktów homologicznych, losuje się podzbiór par punktów, na podstawie którego wyznaczana jest macierz fundamentalna  $F$ . w stosowanej implementacji jest to podzbiór 8 par. Następnie sprawdzany jest warunek (4.11) dla macierzy fundamentalnej  $F$  dla całego wstępnego zbioru punktów homologicznych. Jeżeli różnice są większe niż przyjęty próg, macierz jest odrzucana. Natomiast jeśli różnice są poniżej progu, macierz jest zapamiętywana. Po wykonaniu przyjętej liczby iteracji, wybierana jest najlepsza macierz fundamentalna. w stosowanej implementacji przyjęto 1000 iteracji. Ostatecznie, dla najlepszej macierzy fundamentalnej, sprawdzany jest warunek (4.11) i przyjmując próg poprawności, wybierany jest ostateczny podzbiór poprawnych punktów homologicznych. w ten sposób można odfiltrować nawet do 99 % błędnych sparowań. Na rys. 4.23 widać błędne korelacje pozostałe pomimo przeprowadzenia filtracji metodą RANSAC.

Końcowy etap procesu stereowizji stanowi triangulacja współrzędnych, która zostanie przedstawiona w następnym podrozdziale.

#### 4.17. Triangulacja chmury punktów

Posiadając elementy orientacji wewnętrznej i zewnętrznej głowicy stereowizyjnej oraz macierz dysparycji  $D$ , możemy przystąpić do obliczenia współrzędnych przedmiotowych.



Rys. 4.25. Układ kardynalny głowicy stereowizyjnej - rzut poziomy.

Zgodnie z [72] oraz z rys. 4.25 współrzędne obliczymy ze wzorów:



$$\begin{aligned}X(m, n) &= \frac{T \cdot n}{D(m, n)}, \\Y(m, n) &= \frac{T \cdot m}{D(m, n)}, \\Z(m, n) &= \frac{T \cdot f}{D(m, n)}.\end{aligned}\tag{4.33}$$

Powyższe wzory opierają się na podobieństwie trójkątów. Schemat na rys. 4.25 przedstawia przekrój poziomy układu kardynalnego z zaznaczonymi jego kluczowymi elementami.

Ponieważ powyższe równania zostały wyprowadzone z podobieństwa trójkątów, metoda nazywana została triangulacją.

W następnym rozdziale omówione zostaną zestawy obrazów, na których będą testowane zaproponowane metody poszukiwania odpowiedniości.

## 5. MATERIAŁ BADAWCZY

Opracowane metody poszukiwania odpowiedniości zostaną przetestowane na trzech grupach obrazów:

- **Pierwszą grupę** stanowią obrazy elementarnych brył, znanych ze stereometrii, pełniących również kluczową rolę w procesie tworzenia metod i ich weryfikacji.
- Ponieważ metody parowania obrazów mają funkcjonować w domenie zobrazowań medycznych, **drugą grupę** stanowią zobrazowania medyczne. Jako reprezentatywne zobrazowania medyczne, wybrano obrazy naczyń podskórnych wykrytych iluminatorem naczyniowym. o ile przydatność obrazów brył stereometrycznych była uzasadniona na najbardziej pierwotnym etapie prac nad metodami parowania obrazów, tak zobrazowania naczyń są naturalnym wyborem w następnej fazie testów, gdzie konieczne jest potwierdzenie funkcjonalności metod w dziedzinie medycznej.
- **Trzecią grupę** stanowią obrazy testowe udostępnione przez laboratorium stereowizji Middlebury College. Wybór powszechnie stosowanych obrazów porównawczych, jest uzasadniony nie tyle z punktu widzenia potwierdzenia funkcjonalności nowych metod parowania, które niniejsza praca ma proponować, co potwierdzenia prawidłowości wyboru istniejących metod o wysokiej skuteczności działania, dla celów analizy porównawczej.

W celu przeprowadzenia analizy ilościowej otrzymanych wyników poszukiwania odpowiedniości dla każdej z trzech grup obrazów należy dysponować **dysparycjami wzorcowymi**. Dla pierwszej i drugiej grupy obrazów dysparycje wzorcowe zostaną specjalnie opracowane na potrzeby niniejszej pracy. Dysparycje wzorcowe dla trzeciej grupy obrazów opracowane zostały przez laboratorium stereowizji Middlebury College.

W następnych podrozdziałach przedstawione zostaną szczegółowo wymienione wcześniej trzy grupy obrazów, tj. bryły, naczynia oraz obrazy Middlebury.

## 5.1. Modele stereometryczne

Wybór brył stereometrycznych przedstawionych na rys. 5.1, z charakterystycznymi, wyrażeniami zaznaczonymi krawędziami, podyktowany został potrzebą posiadania wzorca, który podczas testów nowych algorytmów, można by poddać szybkiej i nieskomplikowanej analizie i weryfikacji. Ze względu na elementarność kształtu wybrane bryły charakteryzują się również przewidywalnym wynikiem binaryzacji oraz możliwością szybkiego sporządzenia wzorcowych tablic dysparycji. Stworzyło to możliwość, by podczas testów algorytmów, sprawnie, wiersz po wierszu, weryfikować ich działanie oraz lokalizować i usuwać błędy w działaniu. Nie byłoby to możliwe w przypadku obrazów złożonych, zarówno pod względem szczegółów zawartości sceny, jak i pod względem radiometrycznym.

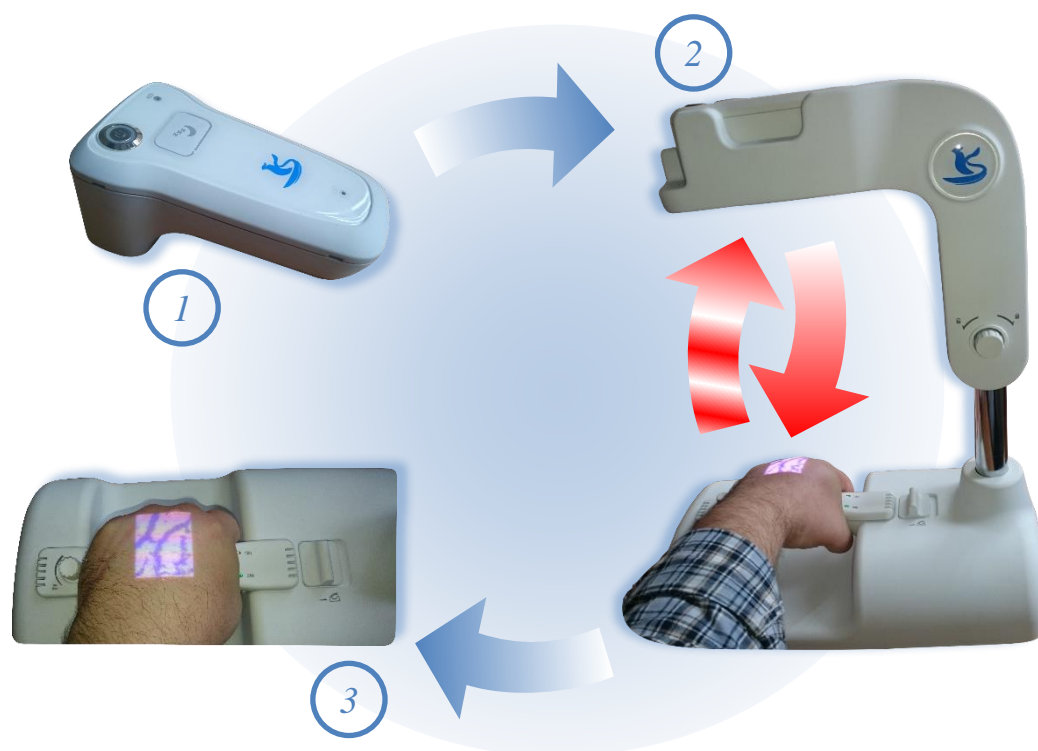


Rys. 5.1. Modele brył stereometrycznych: a) czworościan foremny (Pir); b) sześcian foremny (Cube).

Modele brył stereometrycznych przedstawione na rys. 5.1, tj. sześcian foremny (Cube) i czworościan foremny (Pir) zostały wykonane z kartonu. Wprowadzone krótkie nazwy ułatwiają opisywanie wyników i komponują się z krótkimi nazwami innych obrazów używanych w testach, pochodzących z biblioteki Middlebury. „Pir” to skrót od „Piramida”, potocznego określenia czworościanu foremnego, stosowanego we wcześniejszych publikacjach naukowych [73, 74]. Natomiast „Cube” to angielskojęzyczny odpowiednik sześcianu foremnego. Obrazy zostały skadrowane i przeskalowane do rozdzielczości  $M \times N$  wynoszącej  $200 \times 163$  piksele dla sześcianu i  $200 \times 114$  piksele dla czworościanu. Wymiary obrazów zostały tak ustalone, by zachować jak największe wyraźne obrazy brył, przy jednoczesnym zminimalizowaniu czasu poświęconego na ręczne tworzenie wzorcowej mapy dysparycji  $D_T$ .

## 5.2. Obrazy naczyń podskórnych

Wybór następnej grupy obrazów testowych został podyktowany przeznaczeniem opracowywanych metod poszukiwania odpowiedniości, które mają mieć zastosowanie w medycynie. Ponieważ proponowane metody parowania obrazów mają przyczynić się do stworzenia modelu przestrzennego naczyń podskórnych, zobrazenia powstałe dzięki iluminatorowi naczyniowemu są reprezentatywnym i celowym wyborem. Stanowią one bardzo dobry przykład zobrażeń, z których w drodze segmentacji zostaną wyeliminowane powierzchnie homogeniczne, a następnie wyodrębnione jedynie osie naczyń. w ten sposób następuje minimalizacja liczby elementów, które należy poddać korelacji, co z punktu widzenia analizy złożoności obliczeniowej, na pewno przyspieszy działanie algorytmów. Jednocześnie nastąpi uproszczenie i uporządkowanie przedstawionej na obrazie sceny do jedynie niezbędnych obiektów potrzebnych do podjęcia decyzji przez personel medyczny.



Rys. 5.2. Sureview™ Vein Illumination System ZD-JM-260: 1) analizator z systemem nadawczo odbiorczym podczerwieni wraz z projektorem; 2) system składa się z uchwytu z podstawą, głównego modułu analizatora, oraz dodatkowego emitera podczerwieni, pełniącego rolę wzmacniacza; 3) urządzenie po analizie obrazu z kamery podczerwieni, rzutuje obraz naczyń na powierzchnię skóry.

Obrazy pozyskano zgodnie z Deklaracją Helsińską dokonując rejestracji obrazów na ochotnikach w Katedrze Podstawowych Nauk Biomedycznych na Śląskim Uniwersytecie

Medycznym. Należy zwrócić uwagę, iż w ramach niniejszej rozprawy nie wykonywano żadnych doświadczeń, czy testów z udziałem ludzi. Pacjenci dobrowolnie zgodzili się na rejestrację, a pozyskane dane zostały pozbawione cech umożliwiających personalne rozpoznanie.

Zarejestrowane obrazy testowe przedstawiają projekcję uzyskaną iluminatorem naczyniowym Sureview™ ZD-JM-260 firmy ZD Medical Inc przedstawionym na rys. 5.2. Urządzenie jest stosowane do lokalizacji naczyń podskórnych i jest szczególnie użyteczne podczas infuzji dożyłnej, pobieraniu krwi oraz obliteracji. Oparte na medycznym zimnym świetle jest bezpieczne dla operatora i dla pacjenta. Iluminator wysyła wiązkę bliskiej podczerwieni (od 750 nm do 980 nm), która jest pochłaniana przez krew i odbijana przez tkankę miękką. Odbity obraz jest rejestrowany, analizowany i w czasie rzeczywistym wyświetlany na skórze. Natężenie oświetlenia z projektora można ustawić w zakresie od 300 do 1000 luksów. Efektywna odległość projekcji to  $30\text{ cm} \pm 1\text{ cm}$  przy wymiarach rzutowanego obrazu  $\sim 3,5 \times 3,5\text{ cm}$ .

Tab. 5.1. przedstawia podstawową metrykę pacjentów, na których dokonano rejestracji obrazów naczyń podskórnych.

Tab. 5.1. Metryka pacjentów.

Wiek			Rasa <sup>1</sup>	Stan zdrowia		BMI		
0 - 18	19 - 45	46 - 75		zdrowi	chorzy	< 18,5	18,5 – 24,99	≥ 25
0	3	1	Kaukaska	3	1	0	3	1

Wykonano 100 zobrazowań, a następnie wyselekcjonowano 10 charakterystycznych, powtarzających się układów naczyń. Opracowywane metody segmentacji i parowania obrazów nie działają w oparciu o założenia z dziedziny probabilistyki i dlatego potwierdza się jedynie skuteczność metod dla charakterystycznych powtarzalnych układów naczyń. z uwagi na deterministyczny charakter metod, przyjmuje się powtarzalność wyników dla zobrazowań naczyń o zbliżonych cechach.

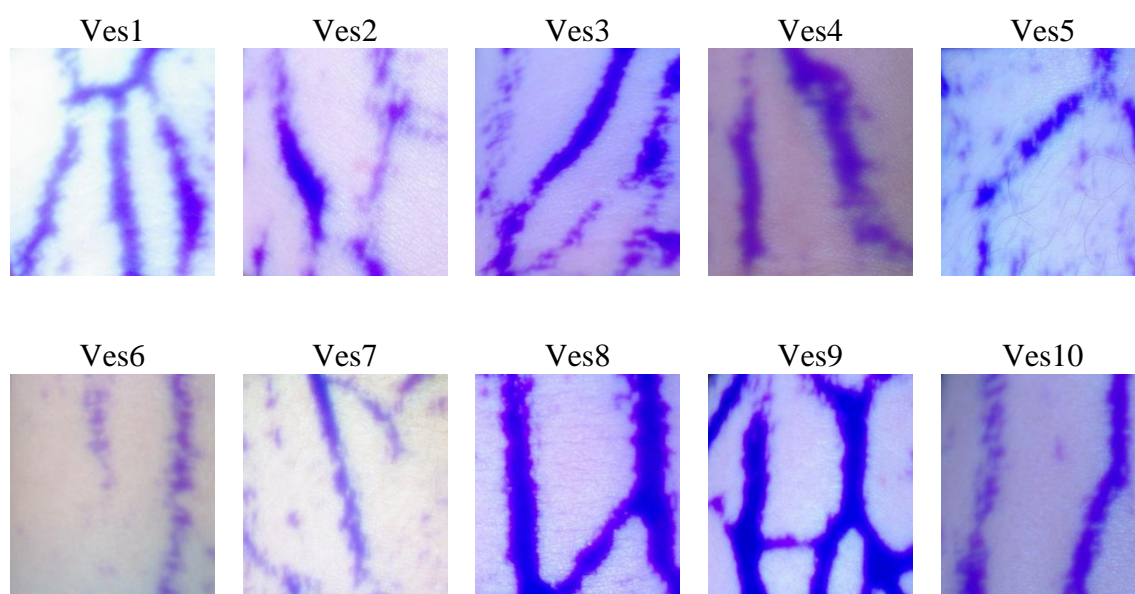
Z punktu widzenia budowy i działania algorytmów parowania obrazów, charakterystyczne będą obrazy o:

<sup>1</sup> Według koncepcji typologii rasowej Hooton'a.

- tej samej liczbie naczyń,
- zbliżonej lokalizacji i orientacji naczyń,
- zbliżonym śladzie przebiegu, krzyżowania się, zbiegania i rozbiegania się naczyń.

Wyselekcjonowane w ten sposób obrazy przedstawione zostały na rys. 5.3.

Przedstawiają one układ naczyń na nadgarstku, przedramieniu, podudziu i grzbietowej części dłoni. Zobrazowania zostały skadrowane do wysokości  $M=300$  punktów, szerokości  $N \in (300, 400)$  odpowiedniej dla rozpiętości naczyń oraz dysparycji wzorcowej  $D_T$  równej 20 lub 30 punktów. Treść obrazów nie była modyfikowana w żaden sposób, tj. żadne elementy nie były z obrazów usuwane lub dodawane.



Rys. 5.3. Testowane obrazy z iluminatora naczyńowego: Ves1, Ves2, Ves3, Ves4 obrazy z nadgarstka; Ves5 obraz z podudzia; Ves6, Ves7 obrazy z przedramienia; Ves8, Ves9, Ves10 obrazy grzbietowej części dłoni.

W następnym podrozdziale przedstawione zostaną obrazy rzeczywiste udostępnione przez Laboratorium Stereowizji z Middlebury College.

### 5.3. Rzeczywiste obrazy z Middlebury

Zestaw 21 obrazów z biblioteki Middlebury College, które są przedstawione na rys. 5.4 stanowi bazę porównawczą dla wszystkich nowych metod poszukiwania stereokorespondencji. Wykorzystywany w testach zestaw obrazów został pozyskany latem 2006 roku przez zespół



kierowany przez Daniela Scharstein'a [75–77]. Obrazy te posiadają przygotowane przez autorów wzorcowe tablice dysparycji.



Rys. 5.4. Obrazy testowe udostępniane przez Laboratorium Stereowizji Middlebury College. Nazwy podano z jęz. angielskim ze względu na konieczność zachowania oryginalnej pisowni.

Obrazy standaryzowane wykorzystywane w testach posiadają ujednoliconą wysokość  $M = 370$  pikseli i różną szerokość  $N \in (400, 500)$ . Obrazy zawierają problematyczne sceny z punktu widzenia poszukiwania odpowiedniości pomiędzy stereoparami, między innymi zawierające:

- elementy nisko-kontrastowe, np.: Rocks, Wood,
- powierzchnie homogeniczne, np.: Plastic, Lampshade,
- znaczne powierzchnie z drobnym, powtarzalnym wzorcem, np.: Aloe, Cloth,
- połączenie powyższych, np.: Baby, Monopoly.

Biblioteka obrazów z Middlebury doskonale nadaje się do potwierdzenia skuteczności działania algorytmów parowania obrazów NCC, VAR i SGBM z którymi będą porównywane proponowane algorytmy.

W następnym podrozdziale przedstawione zostanie zagadnienie pozyskania dysparycji wzorcowych.

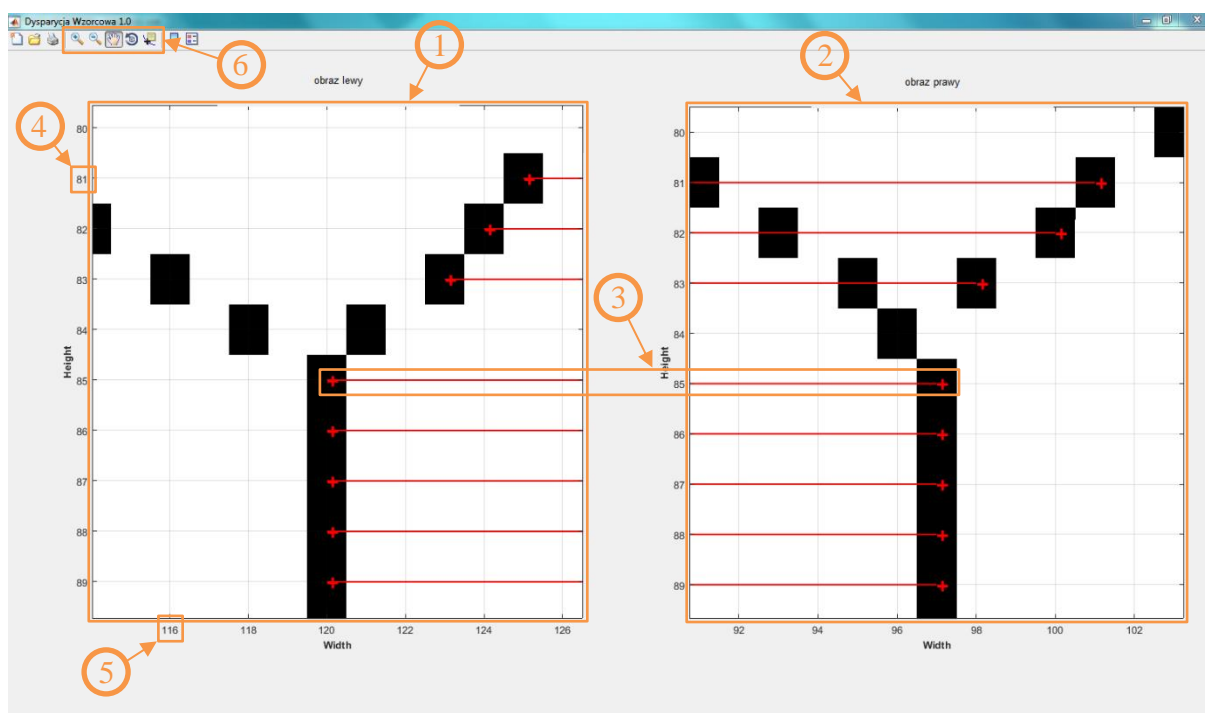
#### 5.4. Macierz dysparycji wzorcowych

Wykonanie analizy ilościowej skuteczności działania metod poszukiwania odpowiedniości wymaga posiadania wzorcowej macierzy dysparycji  $D_T$  do której będzie można się odnieść podczas testowania algorytmu. Dla obrazów kolorowych powszechnie pozyskuje się wzorcowe dysparycje wykorzystując skanery laserowe lub światło strukturalne [75]. Choć pozyskane w ten sposób dysparycje uznaje się za bezbłędne, to są one jednak obarczone błędami systematycznymi i przypadkowymi. Podczas porównywania nowo obliczonych dysparycji do wzorcowych, należy w takim przypadku przyjmować pewien przedział wartości za poprawny. w niniejszej pracy stosuje się przedział 2 pikseli we wszelkich porównaniach, jako implikujący wartość poprawną dysparycji nowo obliczonej.

W większości zobrazowań wykorzystywanych przez autora wielkość rzeczywista piksela zawiera się w przedziale pomiędzy  $70 \mu\text{m}$  a  $950 \mu\text{m}$ , co jest uzależnione od wielkości piksela na płaszczyźnie obrazowej i odległości przedmiotowej. Zobrazowania naczyń i modeli stereometrycznych charakteryzują się wielkością rzeczywistą piksela  $80 \mu\text{m}$ . Przy założeniu zakresu 2 pikseli podczas porównywania nowo obliczonej dysparycji do dysparycji wzorcowej otrzymujemy w wersji pesymistycznej dokładność rzeczywistą  $\pm 170 \mu\text{m}$  w stosunku do dysparycji wzorcowej. Dla większości zastosowań stereowizji w medycynie i inżynierii jest to



wielkość, którą można uznać za dostatecznie małą, by pozostać bez znaczącego wpływu na końcowy wynik opracowania.



Rys. 5.5. Widok okna roboczego aplikacji do tworzenia dysparycji wzorcowych: 1) obszar roboczy na lewym obrazie; 2) obszar roboczy na obrazie prawym; 3) wektor sygnalizujący wcześniej wskazane przez operatora punkty homologiczne; 4) i 5) oznaczenie wiersza i kolumny umożliwiające łatwą lokalizację na obrazach; 6) zestaw ikon umożliwiających powiększanie, pomniejszanie i przemieszczanie się po obrazach.

W zależności od rodzaju zobrazowań dysparycja wzorcowa powstaje w różny sposób. Zobrazowania naczyń przedstawiają rodzaj płytkiej sceny, którą dodatkowo można uznać za płaską. w takim wypadku wystarczy zarejestrować jedno zobrazowanie. Przyjmując a priori dysparycję wzorcową  $D_T$  równą 20 lub 30 pikseli ze zobrazowania są kadrowane dwa obrazy tworzące stereoparę przez horyzontalne przesunięcie jednego obrazu względem drugiego o wartość dysparycji. Zaznaczyć należy jeszcze raz, iż metoda taka jest słuszna tylko w wypadku płaskiej sceny, gdzie płaszczyzna sceny jest równoległa do płaszczyzny obrazowej.

Do określenia dysparycji wzorcowej dla modeli brył stereometrycznych wykorzystano własną aplikację przedstawioną na rys. 5.5, w której wskazuje się ręcznie piksel na obrazie lewym i odpowiadający mu piksel na obrazie prawym. Jest to metoda pracochłonna i z tego powodu jest ona możliwa do wdrożenia tylko dla obrazów binarnych, przedstawiających elementarne sceny o niewielkiej liczbie elementów.

W pracy [72] (Kraus) różniczkuje wzór na odległość obrazową. Na tej podstawie obliczono dokładności odległości obrazowej  $\delta_z$  dla typowych wartości odległości obrazowej  $Z$ ,

ogniskowej normalnej  $f$  oraz błędu obliczenia paralaksy  $\delta_D$ , przyjętego jako rozmiar piksela na zobrazowaniu i wynik przedstawiono w tab. 5.2.

Tab. 5.2. Dokładności odległości obrazowej  $\delta_Z$  [mm] w funkcji ogniskowej i odległości obrazowej.

$f$ [mm]	$Z$ [cm]	50	100	150	Urządzenie obrazujące $\delta_D$ [ $\mu$ m]
6		0,21	0,83	1,88	LOGITECH HD Pro Webcam C920 2,5
28		0,11	0,43	0,96	Pentax K2000 6
43		0,08	0,33	0,73	Film 35mm (24×36) 7

Z powyższej tabeli wynika, iż zakres pikseli przyjęty podczas porównywania obliczonych dysparycji  $D$  z dysparycjami wzorcowymi  $D_T$  można by w niektórych przypadkach zwiększyć np.: dla odległości obrazowej 1,5 m i urządzenia Logitech nawet 9 krotnie. Jednak z analizy przedmiotowej literatury wynika, iż przyjęło się porównywać dysparycje z tolerancją 1 lub 2 pikseli.

## 6. OPIS METODY ZNAJDYWANIA ODPOWIEDNIOŚCI PUNKTÓW

### 6.1. Akwizycja obrazów testowych

Sposób akwizycji powinien zawsze zostać dostosowany do warunków ekspozycji. Różnorodność warunków oświetleniowych i konfiguracji przestrzennej obiektu i urządzenia obrazującego wymusza adaptację technik i posiadanego sprzętu. w następnych podrozdziałach przedstawione zostaną parametry urządzeń obrazujących oraz konfiguracja sceny.

#### 6.1.1. Rejestracja brył

Rejestrację obrazów przedstawiających bryły stereometryczne przeprowadzono z wykorzystaniem głowicy stereowizyjnej przedstawionej na rys. 6.1, o znanych parametrach orientacji wewnętrznej i zewnętrznej przy zachowaniu odległości obrazowej 1 m i bazy  $T=50$  cm.

Głowica stereowizyjna została zbudowana w oparciu o dwie kamery internetowe Logitech C920 HD Pro Webcam. Wyposażone są one w moduły CMOS 3MP o wysokości 3,6 mm i szerokości 4,8 mm, co daje przekątną i jednocześnie ogniskową standardową równą 6 mm. Dzięki ekstrapolacji wykonywanej przez sterowniki, można zarejestrować strumień w rozdzielczości  $1920 \times 1080$  pikseli. Rejestrowane jednocześnie obrazy lewy i prawy są zapisywane w formacie JPG, w przestrzeni barw RGB i rozdzielczością radiometryczną 8 bit na kanał.

Kamery są wyposażone w obiektyw Tessar wykonany przez zakłady optyczne Carl Zeiss. Konstrukcja obiektywu Tessar charakteryzuje się bardzo dobrymi właściwościami optycznymi. Jedyną nieusuniętą aberracją jest niewielka dystorsja beczkowata. Wartość ogniskowej jest kontrolowana w sposób programistyczny, za pośrednictwem sterowników.

Konstrukcja głowicy została zrealizowana w oparciu o statyw i poziome ramię firmy Manfrotto, które charakteryzuje się stabilną konstrukcją. Kamery są umocowane na ruchomych wózkach, poruszających się po poziomym ramieniu, które pozwala pozycjonować kamery w odległości od 10 cm do 90 cm od siebie. Poziome ramię można dowolnie obracać wokół osi ramienia. Wysokość pionową ramienia można ustawić od 70 cm do 1,6 m.



Rys. 6.1 Głowica stereowizyjna stosowana w badaniach: 1) Logitech C920 HD Pro; 2) ruchomy wózek; 3) poziome ramię; 4) pionowe ramię statywu.

W następnym podrozdziale przedstawiony zostanie problem rejestracji obrazów naczyń podskórnych.

### 6.1.2. Rejestracja naczyń podskórnych

Akwizycja obrazów naczyń podskórnych była możliwa jedynie za pomocą wysokiej klasy obiektywu o dużej jasności. w tym celu wykorzystano aparat Pentax K2000 z obiektywem 18-55 mm, co daje ekwiwalent ogniskowej obiektywu w odniesieniu do „full frame” 27-82 mm. Aparat posiada matrycę CCD formatu APS-C firmy HOYA o rozmiarze 23,5×15,7 mm, co daje wartość przekątnej 1,5 razy krótszą od standardowej. Matryca 10,2 MP dysponuje rozdzielczością 3872 na 2592 piksele. Głębia kolorów wynosi 8 bitów dla zapisu w JPEG i 12 bitów w zapisie RAW.

Zobrazowania wykonano przy odległości obrazowej 50 cm. Rzutowany na powierzchnię skóry przez projektor obraz naczyń charakteryzuje się względnie dużą jasnością względem

otoczenia. Odpowiednie naświetlenie projekcji naskórnej powodowało niedoświetlenie otoczenia, co ujawniało się na obrazach znacznie ciemniejszym obszarem wokół projekcji.

Przygotowanie stereopary było możliwe dzięki spełnieniu dwóch założeń.

1. Scena jest płytka. Całość kadru wypełniona jest elementem ciała z rzutowaną projekcją,
2. Scena ma płaski charakter. Na niewielkim obszarze projekcji powierzchnia ciała jest płaska.

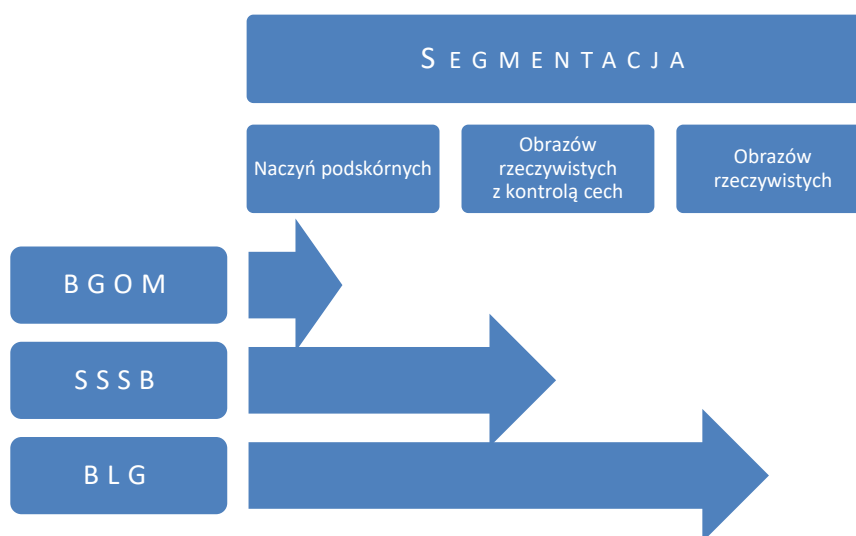
Ze względu na powyższe, macierz dysparycji  $D$  jest jednolita, tzn. wszystkie wartości dysparycji mają zbliżone wartości. Dzięki temu, poprzez odpowiednio wykonane kadrowanie zobrazenia, można je było podzielić na dwa nakładające się zobrazenia, przesunięte poziomo względem siebie, tworzące stereoparę. w tym wypadku wektor przesunięcia odpowiada wartości dysparycji, którą ustala się a priori.

Po akwizycji obrazów następuje etap ich kalibracji i rektyfikacji, a następnie opisane w następnym podrozdziale przetwarzanie wstępne. Etap kalibracji i rektyfikacji jest dobrze opracowany w literaturze i przeprowadzany jest zgodnie ze znanymi już metodami, opisanymi w podrozdziałach Wprowadzenia.

## 6.2. Wstępne przetwarzanie obrazów

Wstępne przetwarzanie obrazów brył, naczyń podskórnych oraz rzeczywistych obrazów zrealizowano za pomocą trzech metod segmentacji, tj. BGOM, SSSB i BLG, zgodnie z rys. 6.2. Ponieważ od jakości segmentacji silnie uzależniony jest wynik parowania, to dla poszczególnych grup obrazów opracowano odmienne metody segmentacji. Jeśli celem jest obliczenie współrzędnych osi naczyń krwionośnych, zastosowanie będzie miała metoda BGOM oparta na binaryzacji globalnej z automatycznym progowaniem. Następnie poprzez operacje morfologiczne pozyskane zostaną osie żył o szerokości 1 piksela.

Gdy celem będzie zbadanie wpływu poszczególnych cech na wynik poszukiwania stereokorespondencji, zastosowanie będzie miała metoda SSSB oparta o binaryzację lokalną, która bazuje na stosunku sumy wartości składowych barwy dla trzech kolejnych pikseli zarówno poprzedniego i następującego wiersza, jak i kolumny. Metoda daje wyniki podobne do detekcji krawędzi metodą Canny'ego, natomiast jej złożoność obliczeniowa jest znacznie mniejsza.



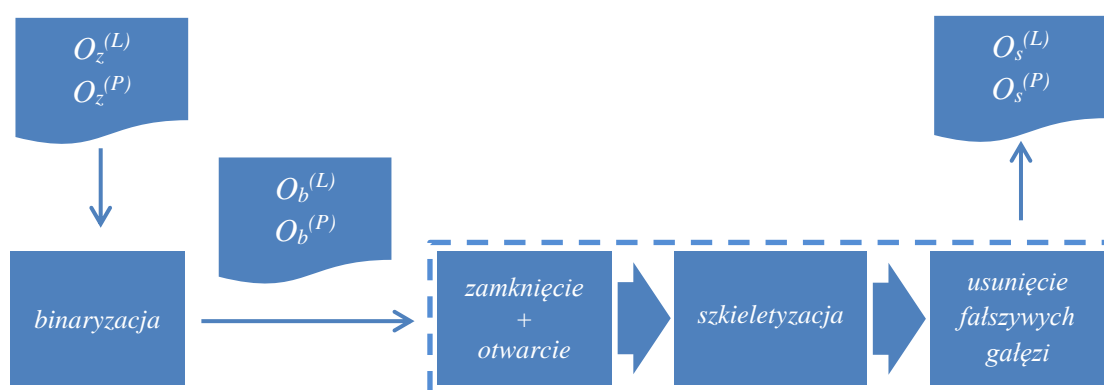
Rys. 6.2 Obszary zastosowań poszczególnych metod segmentacji.

Dla obrazów ogólnych opracowano metodę BLG opartą o binaryzację lokalną i obliczanie gradientu. Metoda ta produkuje obrazy binarne, dla których opracowane metody parowania obrazów dają najlepsze rezultaty dla obrazów rzeczywistych, jednak bez możliwości kontroli cech.

Następnie przedstawione zostaną opracowane metody segmentacji obrazów.

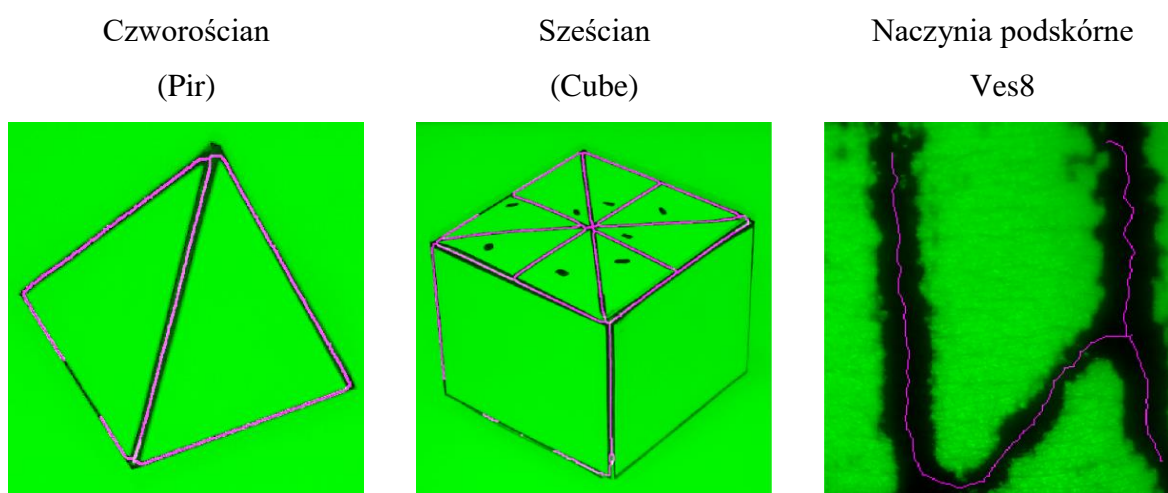
### 6.2.1. Segmentacja metodą BGOM

Zrektyfikowane obrazy lewy  $O_z^{(L)}$  i prawy  $O_z^{(P)}$  binaryzujemy wykorzystując automatyczne progowanie metodą Otsu [78]. Otrzymane w ten sposób obrazy binarne  $O_b^{(L)}$  i  $O_b^{(P)}$  poddajemy procesowi morfologicznego zamknięcia, a następnie otwarcia, usuwając artefakty powstałe w procesie binaryzacji i wygładzając istniejące krawędzie. Na tak przygotowanych obrazach przeprowadzamy operację szkieletyzacji. Po usunięciu bocznych gałęzi otrzymujemy obrazy  $O_s^{(L)}$  i  $O_s^{(P)}$  na których wartości „1” pokrywają się dokładnie z osiami elementów liniowych. Schemat blokowy metody BGOM przedstawiony został na rys. 6.3, zaś wynik poszczególnych etapów segmentacji znajduje się na rys. 6.5. Nazwa metody BGOM jest akronimem od „Binaryzacja globalna z operacjami morfologicznymi”. Wynik szkieletyzacji zależy w znacznej mierze od ustalenia właściwej wartości progowania podczas procesu binaryzacji. Konstruując proces segmentacji, bazując na dotychczasowych doświadczeniach, uznano, iż najlepszą wartość przybliża metoda Otsu, która dla obrazów przedstawiających naczynia, wyraźnie różnicuje linie naczyń od tła.



Rys. 6.3. Proces segmentacji osi naczyń krwionośnych metodą BGOM.

W przypadku modeli figur stereometrycznych zarys krawędzi również został przedstawiony zadawalająco. Wyniki szkieletyzacji nałożone na obrazy binarne przedstawione są na rys. 6.4.

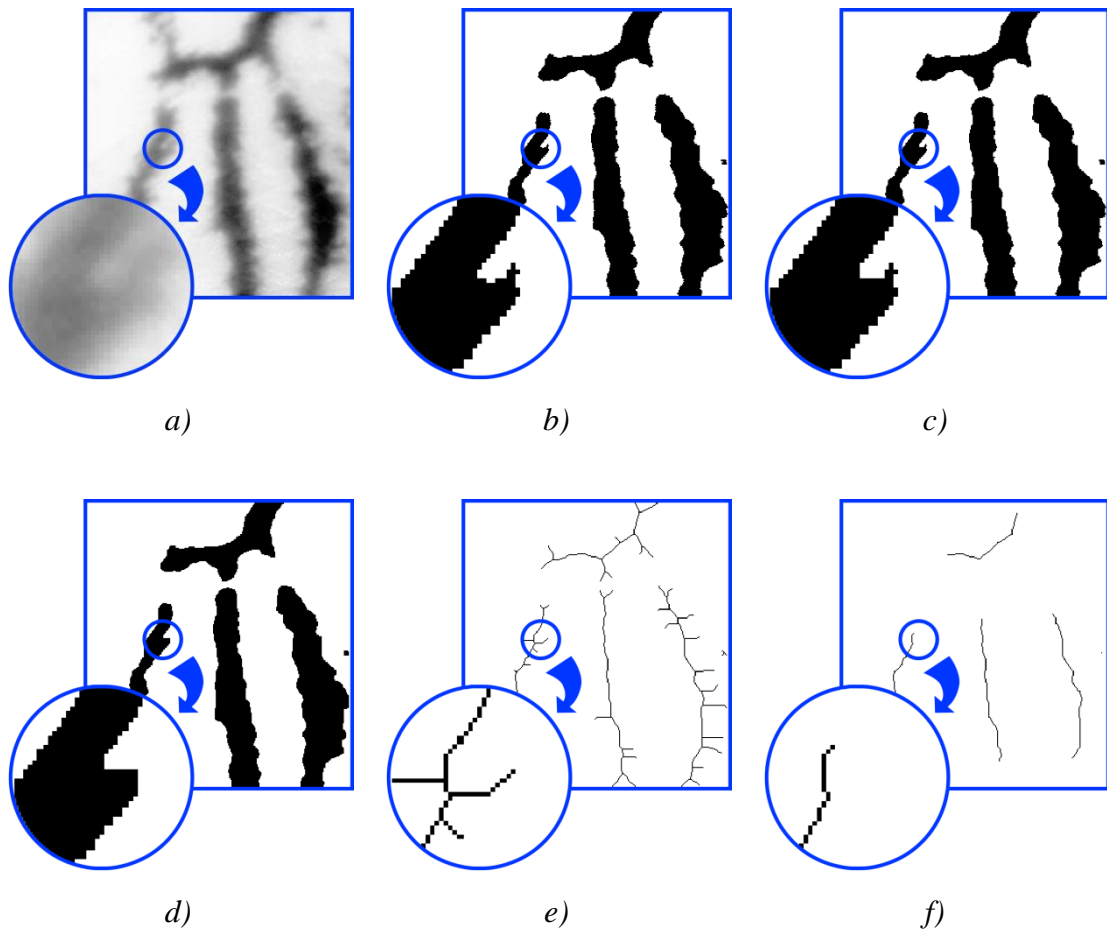


Rys. 6.4. Wysegmentowane osie obiektów liniowych uwidocznione na różowo.

Dla obrazów Middlebury ustalenie progu binaryzacji metodą Otsu niestety się nie sprawdza, dając często obrazy, na których nie można rozpoznać elementów, lub są one silnie zniekształcone po przez zlanie z tłem. Dlatego podczas porównywania wyników parowania dla obrazów Middlebury, dla metod NCC, VAR i SGBM jako obrazy źródłowe przyjęto na wejściu obrazy *RGB* z pominięciem procesu segmentacji. w pozostałych przypadkach, tj. dla wszystkich obrazów naczyń krwionośnych i modeli figur stereometrycznych przyjęto na wejściu wynik segmentacji. Dla nowopropomowanych metod MED, MED-NDD i MED-RGB dla wszystkich obrazów przyjęto na wejściu wynik segmentacji, jako że algorytmy operują na obrazach binarnych.

Zauważyć również należy, iż dla celów infuzji dożylniej, pobrania krwi czy iniekcji nie jest konieczne posiadanie współrzędnych przestrzennych dla 100 % ciągłości naczyń.

Wystarczy wysegmentować jednolitych prostoliniowych odcinków pozwalających przeprowadzić miejscowo właściwą procedurę medyczną.



Rys. 6.5. Poszczególne etapy procesu segmentacji metodą BGOM. Mniejszym niebieskim okręgiem oznaczono obszar źródłowy, którego powiększenie pokazano w większym niebieskim okręgu: *a)* obraz w 256 odcieniach szarości; *b)* wynik binaryzacji; *c)* wynik zamknięcia; *d)* wynik otwarcia; *e)* wynik szkieletyzacji; *f)* obraz po usunięciu bocznych gałęzi.

W następnym podrozdziale przedstawiona zostanie następna metoda segmentacji.

### 6.2.2. Segmentacja metodą SSSB

Inspiracją dla opracowania metody była konieczność kontroli poszczególnych wartości cech obrazu podczas procesu poszukiwania stereokorespondencji. Ponieważ wynik poszukiwania odpowiedniości jest uzależniony od wyników segmentacji, kontrola cech podczas binaryzacji, może być uznana za równoważną, jak podczas procesu poszukiwania



odpowiedniości. Dla cechy, jaką jest barwa, najwygodniejszy sposób polega na przemnażaniu każdego składnika barwy  $j$  przez przyporządkowany jej współczynnik  $k_j$ , zgodnie z ideą

$$S_j(m, n) = \frac{1}{3} \cdot \sum_{m=1}^{m+1} \sum_{j=1}^3 O_z(m, n, j) \cdot k_j \quad (6.1)$$

gdzie  $m \in M$ ,  $n \in N$  – liczba wierszy i kolumn,

$j \in \langle 1, J \rangle$  – kolejne kanały *RGB*, *HSV* lub *Lab*,

$k_j \in \langle 0, 1 \rangle$  – współczynnik danego kanału.

Dzięki temu każdy składnik barwy staje się cechą, którą można dowolnie kontrolować. Metoda zostanie wykorzystana w badaniach wpływu doboru cech na jakość parowania obrazów, przedstawionych w dalszej części rozprawy doktorskiej. Nazwa metody SSSB jest akronimem pochodzącym od „Stosunek sumy składników barwy”.

Metoda przetwarza obraz w dwóch etapach, tworząc macierz  $O_k$ , zawierającą wynik sumowania po kolumnach i  $O_w$ , zawierającą wynik sumowania po wierszach. Pierwsze przejście sumuje wartości barwy dla trzech pikseli w kolumnie  $n$  i  $n+1$ .

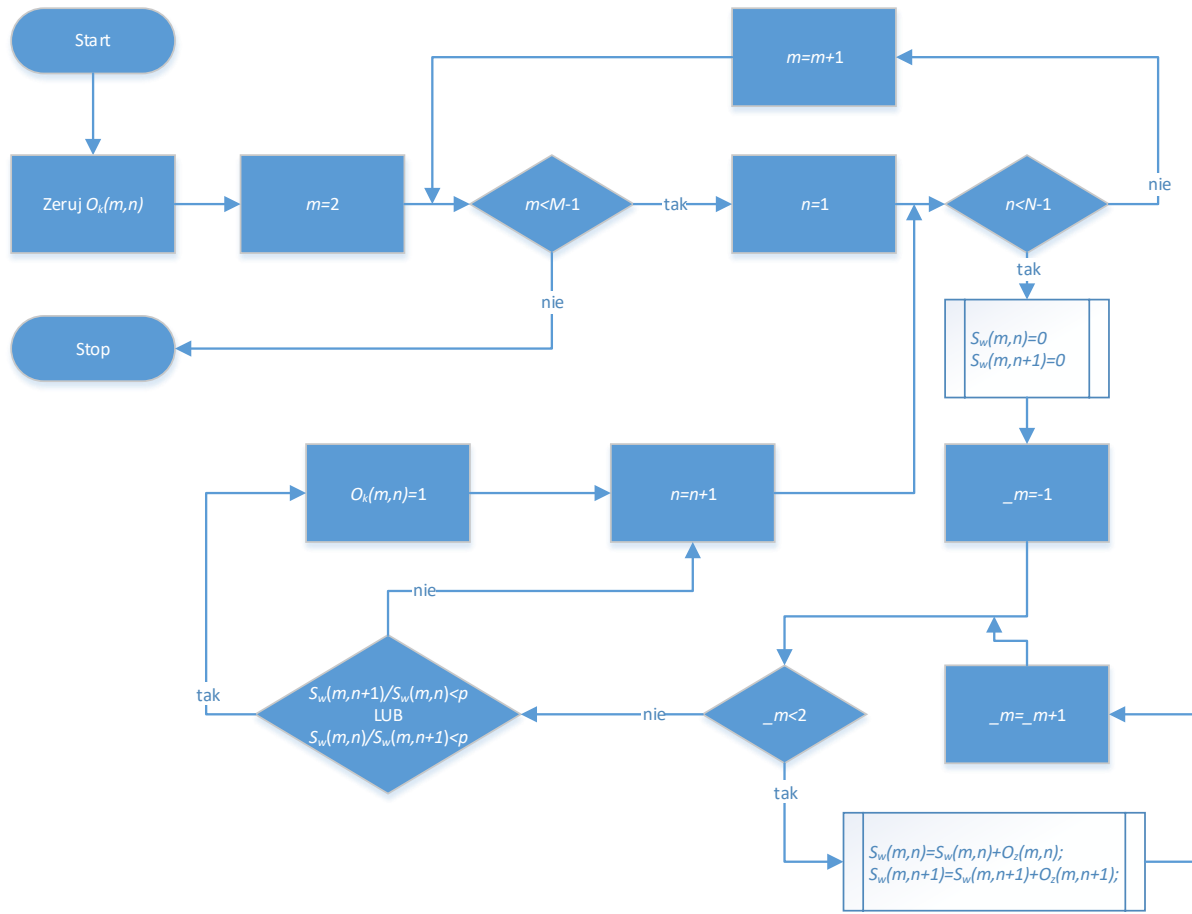
$$S_k(m, n) = \sum_{m=1}^{m+1} \sum_{j=1}^3 O_z(m, n, j) \quad (6.2)$$

$$S_k(m, n+1) = \sum_{m=1}^{m+1} \sum_{j=1}^3 O_z(m, n+1, j) \quad (6.3)$$

Jeśli ich stosunek jest mniejszy od przyjętego progu  $p$  to do macierzy  $O_k$  jest zapisywana jedynka. w przeciwnym wypadku zero.

$$O_k(m, n) = \begin{cases} 1 & \text{dla} \quad \frac{S_k(m, n+1)}{S_k(m, n)} < p \text{ lub } \frac{S_k(m, n)}{S_k(m, n+1)} < p \\ 0 & \text{pozostałe} \end{cases} \quad (6.4)$$

Proces powstania obrazu  $O_k$  przedstawiony został na schemacie blokowym na rys. 6.6.



Rys. 6.6. Schemat blokowy powstawania macierzy  $O_k$  zgodnie z formułą (6.7). Argumentami są macierz  $O_z$  oraz wartość progowa  $p$ .

W drugim przejściu postępujemy analogicznie, sumując wartości barwy dla trzech pikseli w wierszu  $m$  i  $m+1$ .

$$S_w(m, n) = \sum_{n=1}^{n+1} O_z(R_{m,n} + G_{m,n} + B_{m,n}) \quad (6.5)$$

$$S_w(m+1, n) = \sum_{n=1}^{n+1} O_z(R_{m+1,n} + G_{m+1,n} + B_{m+1,n}) \quad (6.6)$$

Macierz  $O_w$  powstaje analogicznie do macierzy  $O_k$

$$O_w(m, n) = \begin{cases} 1 & \text{dla} \quad \frac{S_w(m+1, n)}{S_w(m, n)} < p \text{ lub } \frac{S_w(m, n)}{S_w(m+1, n)} < p \\ 0 & \text{pozostałe} \end{cases} \quad (6.7)$$

Następnie wyniki obydwu przejść są sumowane tworząc obraz binarny  $O_b$ , odpowiednio dla obrazu lewego i prawego, w taki sposób że

$$O_b^{(L)}(m, n) = \begin{cases} 1 & \text{dla } O_k^{(L)}(m, n) + O_w^{(L)}(m, n) > 0 \\ 0 & \text{pozostałe} \end{cases} \quad (6.8)$$

$$O_b^{(P)}(m, n) = \begin{cases} 1 & \text{dla } O_w^{(P)}(m, n) + O_k^{(P)}(m, n) > 0 \\ 0 & \text{pozostałe} \end{cases} \quad (6.9)$$

Otrzymane obrazy binarne  $O_b^{(L)}$  i  $O_b^{(P)}$  stanowią podstawę do obliczania różnicy między sąsiednimi pikselami w poziomie, obliczanej dla każdego wiersza (6.10)(6.11).

$$O_{rw}^{(L)}(m, n) = \begin{cases} 1 & \text{dla } O_b^{(L)}(m, n+1) - O_b^{(L)}(m, n) > 0 \\ 0 & \text{pozostałe} \end{cases} \quad (6.10)$$

$$O_{rw}^{(P)}(m, n) = \begin{cases} 1 & \text{dla } O_b^{(P)}(m, n+1) - O_b^{(P)}(m, n) > 0 \\ 0 & \text{pozostałe} \end{cases} \quad (6.11)$$

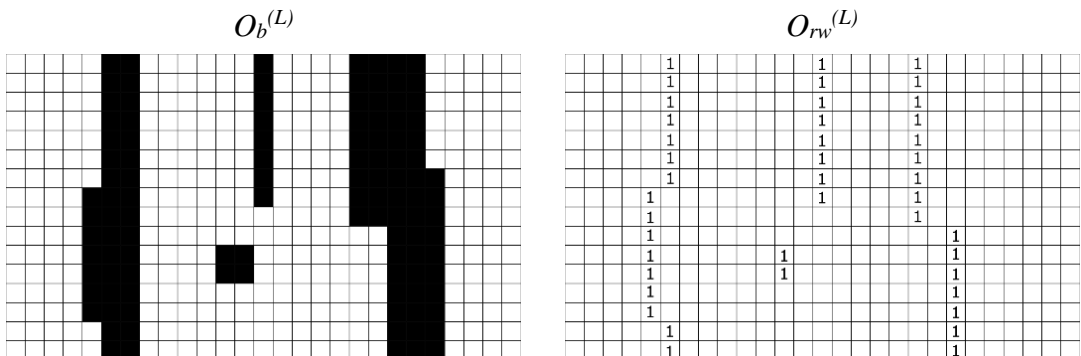
oraz w pionie, obliczanej dla każdej kolumny (6.12)(6.13)

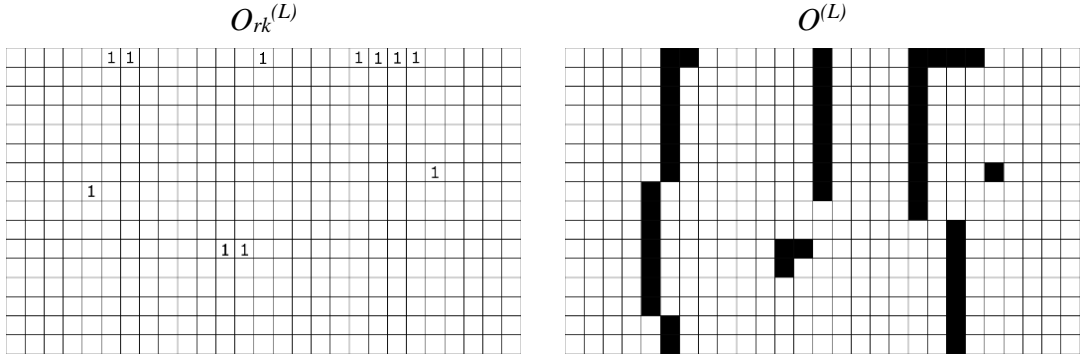
$$O_{rk}^{(L)}(m, n) = \begin{cases} 1 & \text{dla } O_b^{(L)}(m+1, n) - O_b^{(L)}(m, n) > 0 \\ 0 & \text{pozostałe} \end{cases} \quad (6.12)$$

$$O_{rk}^{(P)}(m, n) = \begin{cases} 1 & \text{dla } O_b^{(P)}(m+1, n) - O_b^{(P)}(m, n) > 0 \\ 0 & \text{pozostałe} \end{cases} \quad (6.13)$$

dla  $n \in (1, N-1)$ ,  $m \in (1, M-1)$

Wynikowe obrazy  $O_{rw}^{(L)}$ ,  $O_{rw}^{(P)}$ ,  $O_{rk}^{(L)}$  i  $O_{rk}^{(P)}$  przedstawiają dodatni wynik gradientu, gdzie krawędzie obiektów reprezentowane są przez wartości 1. Patrząc na rys. 6.7, wartości „+1” znajdują się na krawędziach wynikających ze zmiany jasności ciemny – jasny, patrząc od lewej strony.





Rys. 6.7. Binarny obraz  $O_b^{(L)}$ , obraz po zróżnicowaniu  $O_{rw}^{(L)}$  i  $O_{rk}^{(L)}$  oraz obraz finalny  $O^{(L)}$ .

Tworzymy obraz  $O^{(L)}$  i  $O^{(P)}$  dodając do siebie obrazy  $O_{rw}$  i  $O_{rk}$  w taki sposób że

$$O^{(L)}(m,n) = \begin{cases} 1 & \text{dla } O_{rw}^{(L)}(m,n) + O_{rk}^{(L)}(m,n) > 0 \\ 0 & \text{pozostałe} \end{cases} \quad (6.14)$$

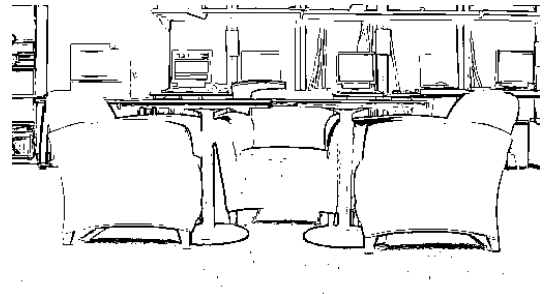
$$O^{(P)}(m,n) = \begin{cases} 1 & \text{dla } O_{rw}^{(P)}(m,n) + O_{rk}^{(P)}(m,n) > 0 \\ 0 & \text{pozostałe} \end{cases} \quad (6.15)$$

dla  $n \in (1, N-1)$

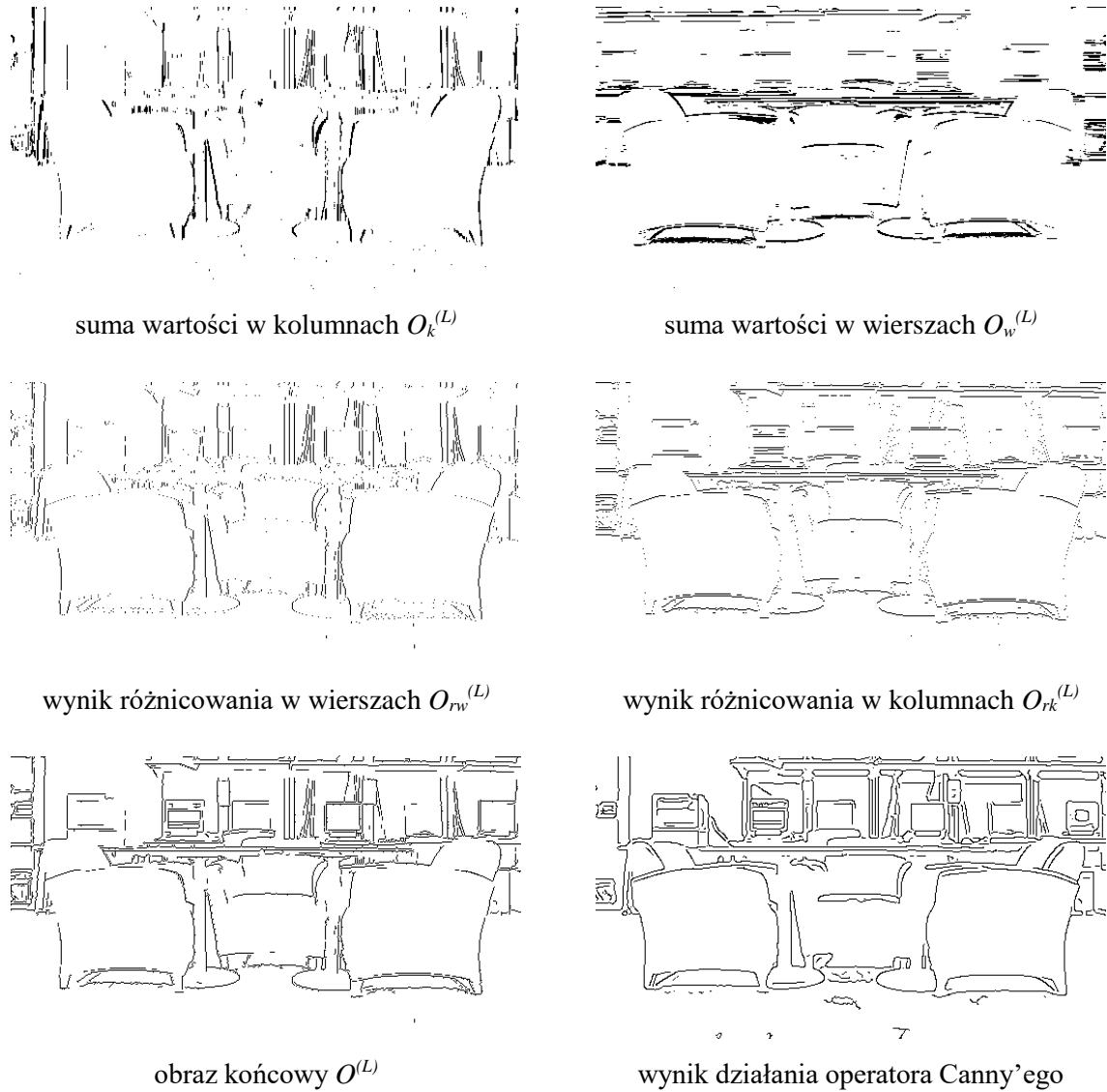
W tak otrzymanych obrazach  $O^{(L)}$  i  $O^{(P)}$  wartości „1” pokrywają się dokładnie z wewnętrznymi krawędziami obiektów o większej intensywności. Obrazy wynikowe dla poszczególnych etapów powstawania obrazu lewego  $O^{(L)}$  zostały przedstawione na rys. 6.8. Ostatni obraz przedstawia wynik działania operatora Canny’ego dla tego samego obrazu. Wynik jest bardzo zbliżony, choć można zaobserwować, iż wynik działania metody SSSB ma mniej szczegółów. Jednak posiada znacznie mniejszą złożoność czasową.



obraz źródłowy *scene\_left.png*



wynikowy obraz binarny  $O_b^{(L)}$



Rys. 6.8. Wyniki poszczególnych etapów segmentacji metodą SSSB.

Następnie przedstawiona metoda segmentacji przedstawia odmienną metodę binaryzacji, opartą również o metodę lokalnej analizy otoczenia.

### 6.2.3. Segmentacja metodą BLG

Do binaryzacji obrazów pochodzących z biblioteki Middlebury wykorzystano metodę binaryzacji lokalnej Sauvola [79]. Próby zbinaryzowania tych obrazów innymi metodami tj. progowaniem globalnym z automatycznym progowaniem metodą Otsu [78], lokalnym progowaniem Bradley'a czy Wolf'a nie przynosiły zadowalających rezultatów, tj. wyniki parowania obciążone były dużymi wartościami błędów parowania obrazów  $\delta$  oraz procentowej

liczby błędnych sparowań  $\delta_D$ . Nazwa metody jest akronimem pochodzącym od „Binaryzacja lokalna z gradientem”.

W wyniku binaryzacji metodą Sauvol’a otrzymano obrazy binarne  $O_b^{(L)}$  i  $O_b^{(P)}$ . Następnie obliczono gradient obrazu uzyskując obrazy  $O^{(L)}$  i  $O^{(P)}$  zgodnie z metodą przedstawioną w poprzednim podrozdziale 6.2.2 od formuły (6.10) do (6.15).

Podczas przetwarzania wstępnego można zdecydować o cechach, które będą podlegać procesowi parowania lub będą pośrednio wpływać na wyniki stereokorespondencji. Problem cech zostanie poruszony w następnym podrozdziale.

### 6.3. Pozyskanie cech charakterystycznych ze stereopar

Ważnym etapem w stereowizji jest wybór cech charakterystycznych. Pozyskane cechy decydują o wyborze metody poszukiwania odpowiedniości pomiędzy obrazami.

W klasycznym podejściu stereokorespondencja porównuje do siebie jasności punktów lub ich otoczenia na obrazie lewym i prawym, by stwierdzić ich odpowiedniość. Podczas poszukiwania odpowiedniości pomiędzy zobrazeniami barwnymi, porównujemy do siebie poszczególne składowe barwy punktów lub otoczenia punktów na stereoparze. Cechą składowych barwy będzie waga  $k$ , która będzie decydować o wpływie poszczególnej składowej na postać obrazu binarnego, a co za tym idzie jakości parowanych obrazów.

Cechy charakterystyczne, należy tu rozumieć jako właściwości atrybutów obrazu, lub właściwości prymitywów z których składa się obraz, bądź też właściwości procesów w wyniku których powstają prymitywy, które są w tym charakterystyczne, że mają wpływ na wynik procesu parowania obrazów. Wyłanianie cech charakterystycznych zależy w dużej mierze od postaci algorytmu poszukiwania odpowiedniości obrazów.

Jeśli algorytm parowania działał by w przestrzeni grafów to cechami mogłyby być odległości radiometryczne pomiędzy sąsiednimi pikselami lub pomiędzy danym pikselem, a wszystkimi pikselami na obrazie.

Podczas parowania prymitywów za cechy charakterystyczne można by uznać np. długość i grubość krawędzi, wielkość skupisk lub pole powierzchni obszarów zamkniętych.

Ponieważ przedstawione w rozprawie doktorskiej algorytmy parowania obrazów funkcjonują w przestrzeni obrazów binarnych, zobrażenia barwne muszą zostać zbinaryzowane. Do binaryzacji zostanie wykorzystany zaproponowany w Rozdziale 6.2.2 algorytm, który umożliwia analizowanie cech składników barwy, o przynależności do

obszarów „zer” lub „jedynek” zdecyduje wartość progowa  $p$ , która stanowi następną ważną cechę obrazów barwnych. Decyduje ona o odpowiedniości obrazu binarnego obrazowi barwnemu.

Transformata Hough’a pozwala znaleźć odcinki proste bazując na zbiorach kolinearnych pikseli w przestrzeni danych. Proste są wyliczane z przestrzeni parametrów transformaty Hough’a. Parametry, które decydują o postaci prostych i co za tym idzie mają wpływ na wynik stereokorespondencji, stanowią cechy charakterystyczne obrazów w przestrzeni Hough’a. Do takich cech zaliczyć można liczbę maksimów lokalnych, wartości progowe zaliczenia maksimum lokalnego, wielkość otoczenia maksimów lokalnych lub ich gęstość, rozmycie maksima lokalnego.

Przejście z obrazów barwnych do obrazów krawędziowych może być zrealizowany operatorem Canny’ego. Parametry, które decydują o ostatecznej postaci krawędzi, mogą być potraktowane jako cechy charakterystyczne obrazów krawędziowych. Złożoność operacji pozwala na wymienienie kilku cech, które decydują o:

- wielkości filtru Gaussa,
- obliczeniu modułu gradientu obrazu,
- pocienianiu (odchudzaniu) krawędzi,
- progowaniu w pętli histerezy.

W następnym podrozdziale przedstawiony zostanie wybór cech znaczących dla przykładowych przestrzeni, w których będą parowane obrazy.

## 6.4. Wybór znaczących cech

Wybranych zostało pięć przestrzeni w których będzie badany wpływ cech na jakość stereokorespondencji:

- przestrzenie barw *RGB*, *HSV* i *Lab*,
- przestrzeń prostych powstałych w wyniku transformacji Hough’a, nazywana dalej przestrzenią Hough’a,
- przestrzeń w postaci obrazu binarnego powstałego w wyniku działania operatora Canny’ego, nazywana dalej przestrzenią Canny’ego.

Dokonując wyboru cech znaczących kierowano się zasadą znacznej zmiany postaci obrazu powodowanej przez niewielką zmianę cechy.

Dla przestrzeni barw *RGB*, *HSV* i *Lab* wybrano odpowiednio jako cechy znaczące:

- wartość progu  $p_{RGB}$ ,  $p_{HSV}$ ,  $p_{Lab}$  który przyjmuje wartości od 0 do 1 ze skokiem co 0,05.
- współczynniki  $k_R$ ,  $k_G$ ,  $k_B$ ,  $k_H$ ,  $k_S$ ,  $k_V$ ,  $k_L$ ,  $k_a$ ,  $k_b$  które przyjmują wartości od 0 do 1 ze skokiem co 0,1, modelujące siłę z jaką dana składowa wpływa na ostateczną barwę.

Proces badania wpływu cech w przestrzeni barw *RGB* przedstawiono w podrozdziale 8.1 i analogicznie dla przestrzeni *HSV* w podrozdziale 8.2 oraz dla przestrzeni *Lab* w podrozdziale 8.3.

Dla przestrzeni Hough'a wybrano jako cechy znaczące:

- minimalną liczbę lokalnych maksimów  $l_{max}$ , które mają zostać wykryte. Wartość  $l_{max}$  odpowiada maksymalnej liczbie prostych, które mają zostać wykryte. Wartość progu jest ustalana w taki sposób, że z wszystkich wykrytych lokalnych maksimów zostaje wybranych  $l_{max}$  największych lokalnych maksimów. Nieodpowiednie ustalenie tej wartości zdecyduje o pominięciu mniejszych lokalnych maksimów, czyli krótszych prostych w przestrzeni danych.
- maksymalna długość pomiędzy dwoma segmentami linii  $d_{max}$ , decydująca o scaleniu tych linii w jedną,
- minimalną długość segmentu linii  $d_{min}$ , która będzie podlegać wykrywaniu. Kształtując tą cechę decydujemy o wartości lokalnego maksimum. Ustalenie za małej wartości może spowodować, iż lokalne maksima będą niewłaściwie wykrywane.

Proces badania wpływu cech w przestrzeni Hough'a przedstawiono w podrozdziale 8.4.

Dla przestrzeni Canny'ego wybrano jako cechy znaczące:

- wartość minimalna progu  $p_{min}$ , dla wykrycia słabych krawędzi oraz wartość maksymalna progu  $p_{max}$ , dla wykrycia wyraźnych krawędzi. Dzięki zastosowaniu dwóch progów operator Canny'ego jest bardzo odporny na szumy. Operator rozpoczyna wykrywanie krawędzi z  $p_{max}$  wykrywając silnie zarysowane krawędzie, następnie przechodzi do  $p_{min}$  i wykrywa słabo zarysowane krawędzie.



- sigmę  $\sigma$  – odchylenie standardowe filtra Gaussowskiego, decydujące o wielkości filtra, czyli o stopniu rozmycia lub wyostrenia obrazu. Gradient obrazu liczony jest jako pochodna filtra Gaussowskiego, gdzie rozmiar filtra jest określany na podstawie  $\sigma$ . Standardowo przyjęło się przyjmować  $\sigma = \sqrt{2}$ .

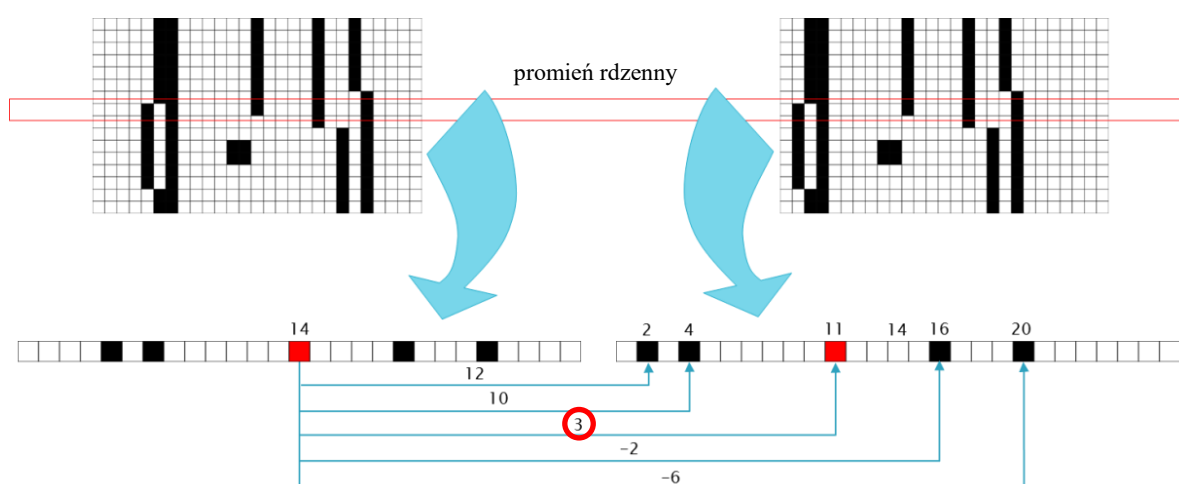
Proces badania wpływu cech w przestrzeni Canny’ego przedstawiono w podrozdziale 8.5.

W następnych podrozdziałach przedstawione zostaną metody parowania obrazów binarnych, opracowanych przez autora niniejszej rozprawy. Następnie przedstawione zostaną wyniki ich działania, również na zobrazowaniach medycznych. w dalszej części pracy przedstawione algorytmy zostaną także użyte do badań wpływu cech na jakość stereokorespondencji.

## 6.5. Metoda poszukiwania stereokorespondencji MED

Pierwszą opracowaną metodą znajdowania odpowiedniości pomiędzy punktami jest metoda MED, której nazwa to akronim angielskiej nazwy *Minimum Element Distance*, tłumaczonej jako – minimalna odległość pomiędzy elementami.

Metoda bazuje na założeniu, że właściwą wartością dysparycji dla zrektyfikowanych obrazów binarnych jest minimalna odległość pomiędzy elementami danego wiersza  $m$  lewego i prawego obrazu [73], co zilustrowano na rys. 6.9.



Rys. 6.9. Wizualizacja procesu parowania obrazów dla algorytmu MED. Właściwa wartość dysparycji oznaczona została czerwonym okręgiem.

Dysponując odpowiednio przygotowanymi obrazami binarnymi możemy przystąpić do skonstruowania wstępnej macierzy odległości lokalnych  $O_o^{(L)}$  i  $O_o^{(P)}$  zawierających informację o odległości od pierwszej kolumny wartości „1” w poszczególnych wierszach zgodnie z poniższymi wzorami

$$O_o^{(L)}(m, n) = \begin{cases} n & \text{dla } O_s^{(L)}(m, n) = 1 \\ 0 & \text{dla } O_s^{(L)}(m, n) = 0 \end{cases} \quad (6.16)$$

$$O_o^{(P)}(m, n) = \begin{cases} n & \text{dla } O_s^{(P)}(m, n) = 1 \\ 0 & \text{dla } O_s^{(P)}(m, n) = 0 \end{cases} \quad (6.17)$$

W celu obliczenia odległości każdego elementu z każdym w odpowiadających sobie wierszach odpowiednio modyfikujemy macierze  $O_o^{(L)}$  i  $O_o^{(P)}$  w taki sposób, że elementy zerowe w każdym wierszu usuwamy, zgodnie z rys. 6.10.

		$O_s^{(L)}$										$O_s^{(P)}$												
		$N$										$N$												
		1	2	3	4	5	6	7	8			1	2	3	4	5	6	7	8					
a)	$M$	1	0	0	1	0	1	0	0	1	$M$	1	0	1	0	1	0	0	1	0				
		2	0	0	0	1	0	0	1	0		2	0	0	1	0	0	1	0	0				
		3	1	0	1	0	1	0	1	0		3	0	1	0	1	0	1	0	0				
		4	0	0	0	0	0	1	0	0		4	0	0	0	0	1	0	0	0				
		5	0	0	0	0	0	0	0	0		5	0	0	0	0	0	0	0	0				
		$O_o^{(L)}$										$O_o^{(P)}$												
		$N$										$N$												
		1	2	3	4	5	6	7	8			1	2	3	4	5	6	7	8					
b)	$M$	1	0	0	3	0	5	0	0	8	$M$	1	0	2	0	4	0	0	7	0				
		2	0	0	0	4	0	0	7	0		2	0	0	3	0	0	6	0	0				
		3	1	0	3	0	5	0	7	0		3	0	2	0	4	0	6	0	0				
		4	0	0	0	0	0	6	0	0		4	0	0	0	0	5	0	0	0				
		5	0	0	0	0	0	0	0	0		5	0	0	0	0	0	0	0	8				
		$O_f^{(L)}$				$O_f^{(P)}$						$K$												
		$I$										1	2	3										
		1	2	3	4							1	2	3										
c)	$M$	1	3	5	8	0							1	2	4	7								
		2	4	7	0	0							2	3	6	0								
		3	1	3	5	7							3	2	4	6								
		4	6	0	0	0							4	5	0	0								
		5	0	0	0	0							5	8	0	0								

Rys. 6.10. Poszczególne etapy budowy macierzy  $O_f^{(L)}$  i  $O_f^{(P)}$ : a) macierz zero-jedynkowa; b) wstępna macierz odległości lokalnych; c) finalna macierz odległości lokalnych.

Następnie dopełniamy każdy wiersz na końcu zerami do zrównania się z najdłuższym niezerowym wierszem tworząc macierz finalną odległości lokalnych  $O_f^{(L)}$  i  $O_f^{(P)}$ . Występujące na rys. 6.10 wielkości  $I, K$  to odpowiednio liczba kolumn macierzy  $O_f^{(L)}$  i  $O_f^{(P)}$  oraz jednocześnie

$$I = \max_m \left( \sum_{n=1}^N O_s^{(L)}(m, n) \right) \quad (6.18)$$

$$K = \max_m \left( \sum_{n=1}^N O_s^{(P)}(m, n) \right). \quad (6.19)$$

Następnie obliczana jest macierz odległości w taki sposób, że różnica odległości pomiędzy punktem na lewym obrazie, a wszystkimi punktami na prawym obrazie w danym wierszu jest równa

$$O_{odl}(m, i, k) = O_f^{(L)'}(m, i, k) - O_f^{(P)'}(m, i, k) \quad (6.20)$$

gdzie

$$O_f^{(L)'}(m, i, k) = [1, 1, \dots, 1]_{1 \times K}^T \cdot O_f^{(L)}(m, i) \quad (6.21)$$

$$O_f^{(P)'}(m, i, k) = O_f^{(P)}(m, k)^T \cdot [1, 1, \dots, 1]_{1 \times I}. \quad (6.22)$$

Przyjmujemy, że wstępna macierz dysparycji  $D_m$  odpowiada minimalnej odległości obliczonej pomiędzy elementem na lewym obrazie, a wszystkimi elementami na prawym obrazie dla danego wiersza  $m$ , zgodnie z

$$D_m(m, i) = \min_k (O_{odl}(m, i, k)). \quad (6.23)$$

Do dalszych obliczeń potrzebna jest nam macierz dysparycji  $D$  o rozmiarze zgodnym z rozmiarem obrazów źródłowych  $M \times N$ , którą obliczymy zgodnie z następującym wzorem

$$D(m, n) = \begin{cases} D_m(m, i) & \text{dla } O_s^{(L)}(m, n) \neq 0 \\ 0 & \text{pozostałe} \end{cases} \quad (6.24)$$

dla

$$m \in (1, M), n = O_s^{(L)}(m, i), i \in (1, I).$$

Dysponując macierzą dysparycji  $D$  można przystąpić do wyliczenia współrzędnych terenowych chmury punktów zgodnie z (4.33).

Algorytm realizujący metodę MED przedstawiony został na rys. 6.11. Jedynym argumentem, który należy zapewnić na wejściu jest stereopara w postaci obrazów binarnych  $O_b^{(L)}$  i  $O_b^{(P)}$ . w wyniku działania algorytmu powstaje macierz dysparycji  $D$ .

W następnym podrozdziale zostanie przeanalizowana złożoność obliczeniowa algorytmu.

### 6.5.1. Analiza złożoności obliczeniowej algorytmu

Na rys. 6.11. przedstawiono algorytm realizujący metodę poszukiwania odpowiedniości MED. Analiza złożoności obliczeniowej zostanie przeprowadzona zgodnie z asymptotycznym tempem wzrostu stosując notację dużego  $O$ . Podczas analizy zachowane będą używane w niniejszej pracy oznaczenia, tj.  $m, n, i, k$  oraz  $M, N, I, K$ , opisujące rozmiar macierzy obrazowych.

```

1   $O_o^{(L)}(m,n) = \text{twórzWstępnąMacierzOdległościLokalnych}(O_b^{(L)});$ 
2   $O_o^{(P)}(m,n) = \text{twórzWstępnąMacierzOdległościLokalnych}(O_b^{(P)});$ 
3  for  $m=1:M$  do
4      |  $O_f^{(L)}(m,i) = \text{twórzFinalnyWektorOdległościLokalnych}(O_o^{(L)});$ 
5      |  $O_f^{(P)}(m,k) = \text{twórzFinalnyWektorOdległościLokalnych}(O_o^{(P)});$ 
6      |  $O_f^{(L)'}(m,i,k) = \text{wektorJedynek}(K)^T \cdot O_f^{(L)}(m,i);$ 
7      |  $O_f^{(P)'}(m,i,k) = O_f^{(P)}(m,k)^T \cdot \text{wektorJedynek}(I);$ 
8      |  $O_{odl}(m,i,k) = O_f^{(L)'}(m,i,k) - O_f^{(P)'}(m,i,k);$ 
9      |  $LiczD_m(m,i);$ 
10     | for  $i = 1:I$  do
11         | |  $D(m, O_f^{(L)}(m,i)) = D_m(m,i);$ 
12     | end for
13 end for
```

Rys. 6.11. Algorytm MED. Argumentem są macierze obrazowe binarne. Funkcja zwraca macierz dysparycji  $D$ .

W pierwszych dwóch liniach kodu tworzona jest wstępna macierz odległości lokalnych  $O_o$  zgodnie ze wzorem (6.16) i (6.17). Przyjmując, że operacja zastąpienia wartości 1 przez  $n$

jest wykonywana w podwójnej pętli przechodząc przez wiersze  $m$  i kolumny  $n$  macierzy obrazu, otrzymuje się złożoność obliczeniową  $O(MN)$ .

W wierszu 4 i 5 tworzona jest finalną macierz odległości lokalnych zgodnie ze sposobem przedstawionym na rys. 6.10. Przyjmując, że utworzy się nowy wektor i do niego przepisze wartości większe od zera, to otrzyma się złożoność obliczeniową  $O(N)$ . w wierszu 6 i 7 ulegają modyfikacji finalne wektory odległości lokalnych tak, by można było je odjąć od siebie i utworzyć macierz odległości o rozmiarze w przypadku najmniej korzystnym  $N \times N$ . Jeśli przyjmie się, że wektor jednostkowy tworzy się w  $i$  lub  $K$  operacjach, gdzie maksymalna ich wartość, w przypadku najmniej korzystnym, to  $N$  oraz wymiar finalnego wektora odległości lokalnych w przypadku najbardziej niekorzystnym to  $N$ , to zgodnie z formułami (6.21) i (6.22) otrzyma się złożoność obliczeniową tych dwóch wierszy  $O(N^2)$ .

W linii 9 obliczana jest wstępna macierz dysparycji  $D_m$  zawierająca minimalne wartości dla każdego wiersza macierzy odległości  $O_{odl}$ . w przypadku najmniej korzystnym złożoność obliczeniowa tej operacji będzie wynosić  $O(N^2)$ . w liniach od 10 do 12 w iteracjach o maksymalnej liczbie  $N$  wstawia się wartości do macierzy dysparycji, więc złożoność obliczeniowa wynosi tu  $O(N)$ .

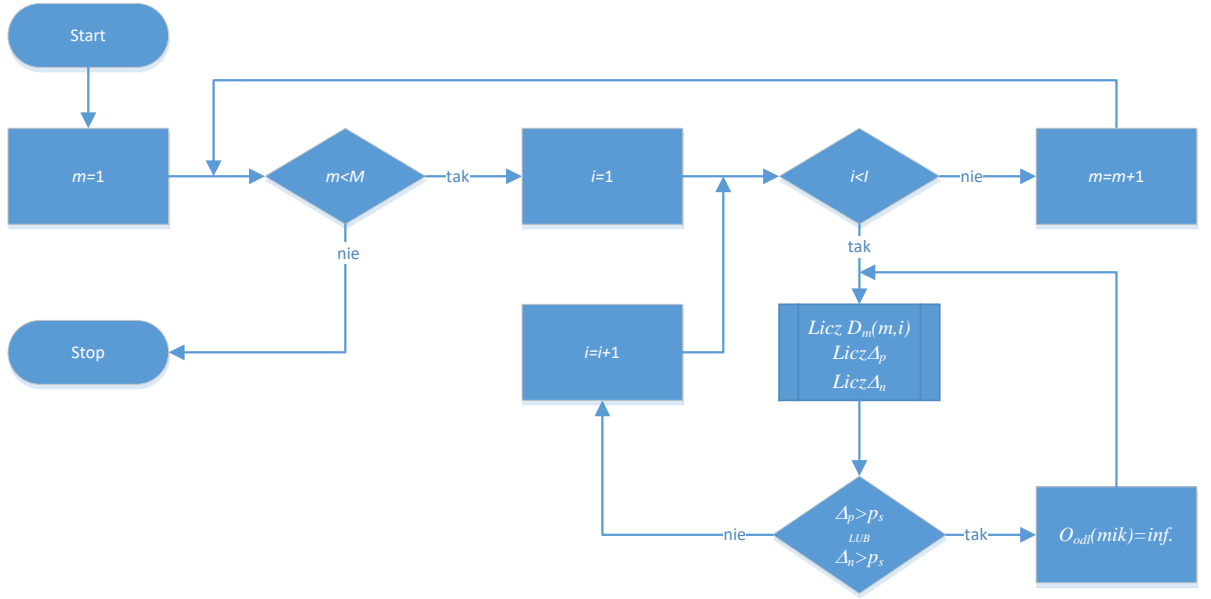
Linie od 4 do 12 iterują  $M$  razy, co daje pesymistyczną złożoność obliczeniową całego algorytmu  $O(MN^2)$ .

## 6.6. Metoda poszukiwania stereokorespondencji MED-NDD

Metoda MED-NDD jest rozwinięciem metody MED o mechanizm eliminacji błędnych sparowań. Dzięki analizie różnic odległości do elementów poprzedzających i następujących po punkcie homologicznym na lewym i prawym obrazie, macierz dysparycji  $D$  zawiera mniej błędnych sparowań, co zostanie dalej wykazane.

Na schemacie blokowy na rys. 6.12 przedstawiono działanie algorytmu weryfikującego macierz wstępnych dysparycji  $D_m$  zgodnie z założeniami metody.

Pierwszy człon nazwy metody jest zaczerpnięty z metody bazowej, drugi zaś to akronim angielskiej nazwy algorytmu *Neighbour Distance Difference*, co należy tłumaczyć jako - Różnica odległości do sąsiada.



Rys. 6.12. Schemat blokowy funkcji weryfikacji macierzy  $D_m$  przez kontrolę odległości do sąsiednich elementów. Argumentem jest macierz  $O_{odl}$ ,  $O_f^{(L)}$ ,  $O_f^{(P)}$ . Funkcja zwraca macierz  $D_m$ .

Przystępując do weryfikacji macierzy  $D_m$  przez kontrolę odległości do sąsiadów dla każdego elementu wstępnej macierzy dysparycji  $D_m$  liczy się odległość do poprzedzającego (6.25) i następującego elementu sąsiadującego (6.26), na obrazie lewym.

$$S_p^{(L)} = O_f^{(L)}(m, i) - O_f^{(L)}(m, i - 1) \quad (6.25)$$

$$S_n^{(L)} = O_f^{(L)}(m, i + 1) - O_f^{(L)}(m, i) \quad (6.26)$$

Analogicznie postępuje się dla obrazu prawego:

$$S_p^{(P)} = O_f^{(P)}(m, k) - O_f^{(P)}(m, k - 1) \quad (6.27)$$

$$S_n^{(P)} = O_f^{(P)}(m, k + 1) - O_f^{(P)}(m, k). \quad (6.28)$$

Odejmuje się odległość do sąsiada poprzedzającego dla rozpatrywanego elementu wstępnej macierzy dysparycji na obrazie prawym od odległości do sąsiada poprzedzającego na obrazie lewym zgodnie ze wzorem

$$\Delta_p = S_p^{(L)} - S_p^{(P)}. \quad (6.29)$$

Następnie analogicznie wylicza się różnicę do sąsiada następującego po rozpatrywanym elemencie, według wzoru

$$\Delta_n = S_n^{(L)} - S_n^{(P)}. \quad (6.30)$$

Jeśli  $\Delta_p$  lub  $\Delta_n$  jest większa niż przyjęty próg  $p_s$ , uznaje się dysparycję za błędną i zastępuje się odpowiadającą jej odległość w macierzy odległości  $O_{odl}$  wartością symbolizującą nieskończoność (6.31), po czym ponownie wyliczamy wstępną macierz dysparycji  $D_m$  zgodnie ze wzorem (6.23).

$$O_{odl}(m, i, k) = \begin{cases} Inf & \text{dla } \Delta_p > p_s \text{ lub } \Delta_n > p_s \\ O_{odl}(m, i, k) & \text{pozostałe} \end{cases} \quad (6.31)$$

W innym wypadku wartość dysparycji uznawana jest za właściwą i pozostawia się ją we wstępnej macierzy dysparycji  $D_m$ . w obliczeniach przyjęto stosować wartość  $p_s=0$ .

Funkcja pracuje na macierzy odległości  $O_{odl}$  oraz finalnych macierzach odległości lokalnych  $O_f$  i zwraca zweryfikowaną wstępną macierz dysparycji  $D_m$ .

Analogicznie, jak dla metody MED, macierz dysparycji  $D$  zostanie obliczona zgodnie z (6.24), a następnie przystąpi się do wyliczenia współrzędnych terenowych chmury punktów zgodnie z (4.33).

W następnym podrozdziale zostanie przeanalizowany algorytm realizujący powyżej przedstawioną metodę MED-NDD pod kątem złożoności obliczeniowej.

#### 6.6.1. Analiza złożoności obliczeniowej algorytmu.

Na rys. 6.13 przedstawiony został algorytm realizujący metodę MED-NDD. Podobnie jak dla poprzednio analizowanej metody, analiza złożoności obliczeniowej zostanie przeprowadzona zgodnie z asymptotycznym tempem wzrostu stosując notację dużego o oraz zostaną zachowane używane w niniejszej pracy oznaczenia, tj.  $m, n, i, k$  oraz  $M, N, I, K$ , opisujące rozmiar macierzy obrazowych.

```

1  for  $m=1:M$  do
2    | for  $i=1:I$  do
3    | |  $done=false$ ;
4    | | do
5    | | |  $LiczD_m(m,i)$ ;
6    | | |  $Licz\Delta_p$ ;
7    | | |  $Licz\Delta_n$ ;
8    | | | if  $(\Delta_p > p_s) \parallel (\Delta_n > p_s)$  then
9    | | | |  $O_{odl}(m,i,k)=inf$ ;
10   | | | else
11   | | | |  $done=true$ ;
12   | | | end if
13   | | while  $\sim done$ 
14   | end for
15 end for

```

Rys. 6.13. Algorytm funkcji weryfikacji macierzy  $D_m$  przez kontrolę odległości do sąsiednich elementów. Argumentem jest macierz  $O_{odl}$ ,  $O_w^{(L)}$ ,  $O_w^{(P)}$ . Funkcja zwraca macierz  $D_m$

Zawartość pętli repetycyjnej zawartej pomiędzy liniami kodu 4 i 13 jest sprzężona z instrukcją warunkową IF i w najbardziej niekorzystnym scenariuszu, wywołana zostanie  $N-1$  razy. w linii 5 liczona jest wstępna macierz dysparycji  $D_m$ . Złożoność obliczeniowa tej operacji w najmniej korzystnym scenariuszu została w rozdziale 6.5.1 ustalona na  $O(N^2)$ . Ponieważ pętla jest wykonywana  $N-1$  razy to otrzymujemy ostateczną złożoność obliczeniową  $O(N^3)$ . Iteracja zawarta w liniach 2–14 może w najmniej korzystnym przypadku, kiedy  $I=N$ , wykonywać się  $N$  razy. Iteracja najbardziej zewnętrzna zawarta w liniach 1–15 zawsze będzie wykona  $M$  razy.

Daje nam to pesymistyczną złożoność obliczeniową całego algorytmu  $O(MN^4)$ .

## 6.7. Metoda poszukiwania stereokorespondencji MED-RGB

Drugą zaproponowaną w niniejszej pracy metodą usunięcia fałszywych sparowań z macierzy dysparycji jest analiza składowych barwy piksela poprzedzającego i następującego punkt homologiczny na obrazie lewym i prawym, co przedstawione jest na rys. 6.14.



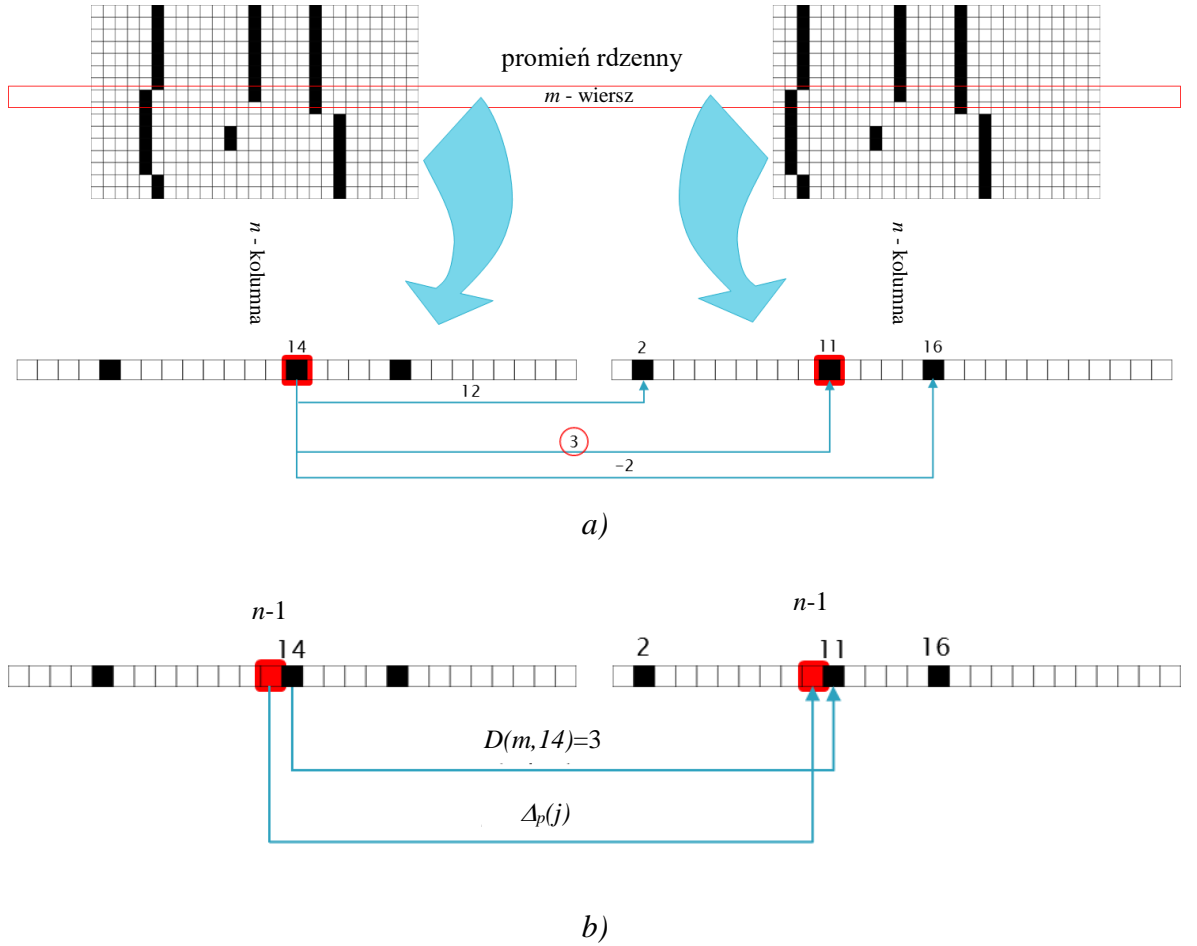
Pierwszy człon nazwy powstał analogicznie jak dla poprzedniej metody. Drugi człon zawdzięcza swoją postać przestrzeni barw, w której, jako pierwszej, metoda była testowana.

W tym celu odejmuje się od siebie składowe  $j$  wektora barwy dla piksela poprzedzającego element homologiczny z wstępnej macierzy dysparycji  $D_m$  na obrazie lewym  $O_z^{(L)}(m, n-1, j)$  i prawym  $O_z^{(P)}(m, n-1, j)$  otrzymując  $\Delta_p(j)$  zgodnie ze wzorem

$$\Delta_p(j) = \left| O_z^{(L)}(m, n-1, j) - O_z^{(P)}(m, n-1, j) \right| \quad (6.32)$$

gdzie

$j = 1, 2, 3$  (kolejne składowe barwy).



Rys. 6.14. Wizualizacja procesu parowania w algorytmie MED-RGB: a) pierwszy etap parowania zgodny z MED; b) – obliczanie odległości radiometrycznej pomiędzy sąsiadami poprzedzającymi punkt homologiczny.

Analogicznie odejmuje się od siebie składowe  $j$  wektora barwy dla piksela następującego po elemencie homologicznym ze wstępnej macierzy dysparycji  $D_m$  na obrazie lewym  $O_z^{(L)}(m, n+1, j)$  i prawym  $O_z^{(P)}(m, n+1, j)$  otrzymując  $\Delta_n(j)$  zgodnie z

$$\Delta_n(j) = \left| O_z^{(L)}(m, n+1, j) - O_z^{(P)}(m, n+1, j) \right|. \quad (6.33)$$

Następnie tworzony jest wektor  $R_p$  i  $R_n$  w taki sposób, że jeśli bezwzględna różnica po danej składowej barwy jest większa od przyjętego progu  $p_k$ , to dla składowej przyjmowana jest wartość 1, w innym wypadku przyjmowana jest 0 zgodnie z

$$R_p(j) = \begin{cases} 1 & \text{dla } \Delta_p(j) > p_k \\ 0 & \Delta_p(j) \leq p_k \end{cases} \quad (6.34)$$

$$R_n(j) = \begin{cases} 1 & \text{dla } \Delta_n(j) > p_k \\ 0 & \Delta_n(j) \leq p_k \end{cases}. \quad (6.35)$$

Jeśli suma elementów wektora  $R_p$  lub  $R_n$  jest większa od progu  $p_s$  to uznaje się dysparycję za błędną i zastępuje się ją w macierzy odległości  $O_{odl}$  wartością symbolizującą nieskończoność  $Inf$ , po czym ponownie liczy się  $D_m$  przeliczając minimalne odległości zgodnie ze wzorem (6.23).

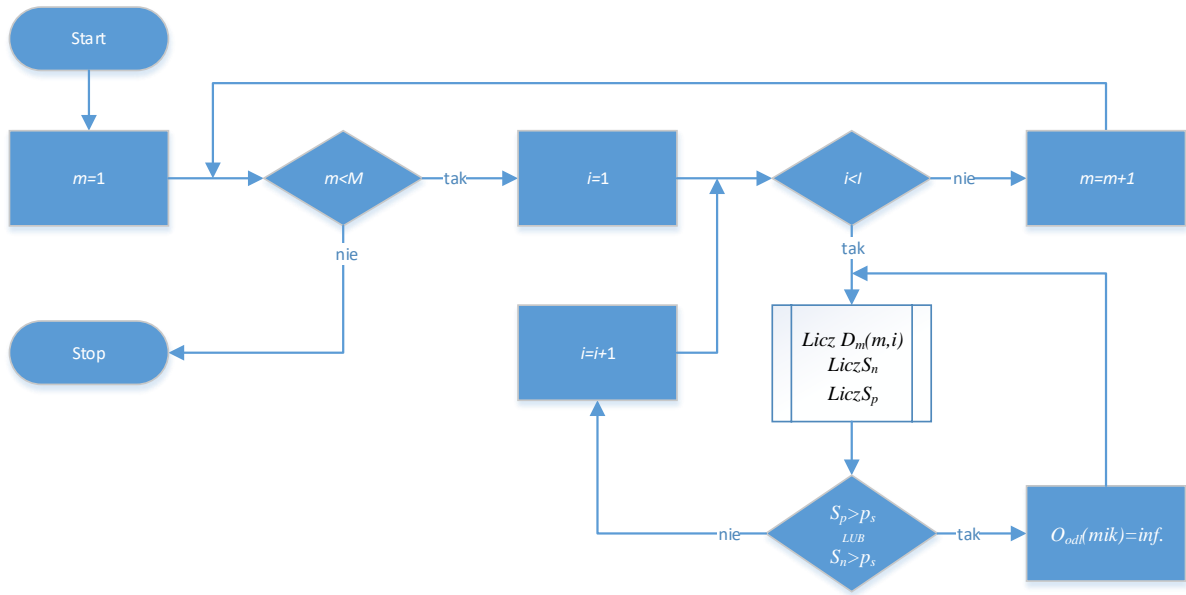
$$O_{odl}(m, i, k) = \begin{cases} Inf & \text{dla } S_p > p_s \text{ lub } S_n > p_s \\ O_{odl}(m, i, k) & \text{pozostałe} \end{cases} \quad (6.36)$$

dla

$$\begin{aligned} S_p &= \sum_{j=1}^3 R_p(j) \\ S_n &= \sum_{j=1}^3 R_n(j) \end{aligned} \quad (6.37)$$

W innym wypadku wartość dysparycji uznawana jest za właściwą i nie dokonuje się zmian w macierzy odległości  $O_{odl}$ . w algorytmie przyjęto wartość  $p_s=0$ .

Schemat blokowy przedstawiony na rys. 6.15 przedstawia działanie algorytmu weryfikującego macierz wstępnych dysparycji  $D_m$  zgodnie z założeniami metody MED-RGB. Analogicznie, jak poprzednia metoda, funkcja pracuje na macierzy odległości  $O_{odl}$  oraz finalnych macierzach odległości lokalnych  $O_f$  i zwraca zweryfikowaną wstępną macierz dysparycji  $D_m$ . Podobnie również, jak dla metody MED i MED-NDD, macierz dysparycji  $D$  zostanie obliczona zgodnie z (6.24), a następnie przystąpi się do wyliczenia współrzędnych terenowych chmury punktów zgodnie z (4.33).



Rys. 6.15. Funkcja weryfikacji macierzy  $D_m$  przez kontrolę różnicy barw sąsiednich pikseli. Argumentem jest macierz  $O_{odl}$ ,  $O_f^{(L)}$ ,  $O_f^{(P)}$ . Funkcja zwraca macierz  $D_m$ .

W następnym podrozdziale przedstawiona zostanie analiza złożoności obliczeniowej algorytmu realizującego powyżej przedstawioną metodę poszukiwania odpowiedniości punktów na obrazach.

### 6.7.1. Analiza złożoności obliczeniowej algorytmu

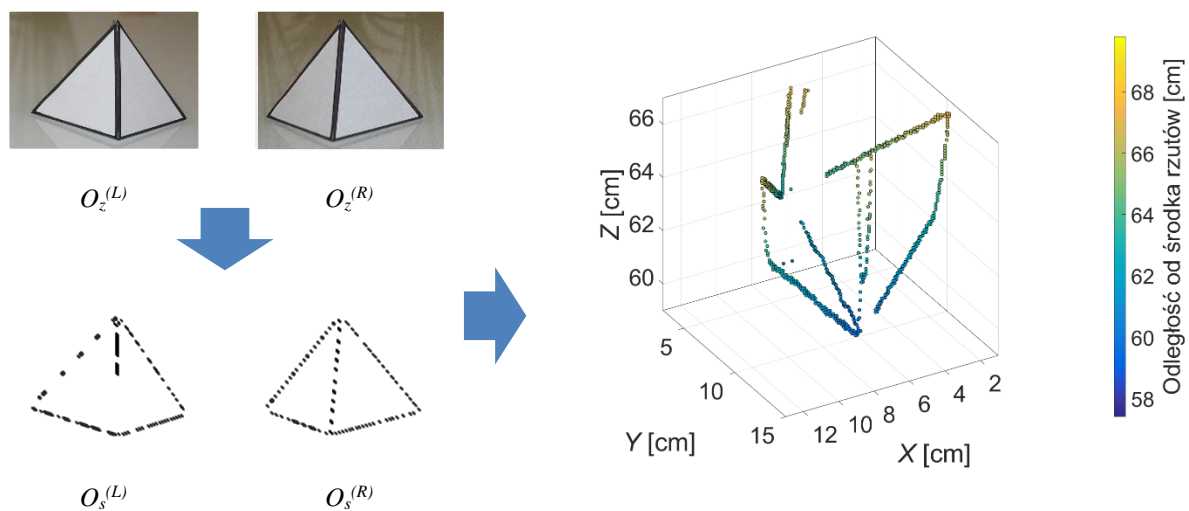
Na rys. 6.16 przedstawiony został algorytm realizujący metodę poszukiwania odpowiedniości MED-RGB pomiędzy punktami obrazów stereowizyjnych. Podobnie jak dla poprzednio analizowanej metody, analiza złożoności obliczeniowej zostanie przeprowadzona zgodnie z asymptotycznym tempem wzrostu stosując notację dużego o oraz zostaną zachowane używane w niniejszej pracy oznaczenia, tj.  $m$ ,  $n$ ,  $i$ ,  $k$  oraz  $M$ ,  $N$ ,  $I$ ,  $K$ , opisujące rozmiar macierzy obrazowych.

```

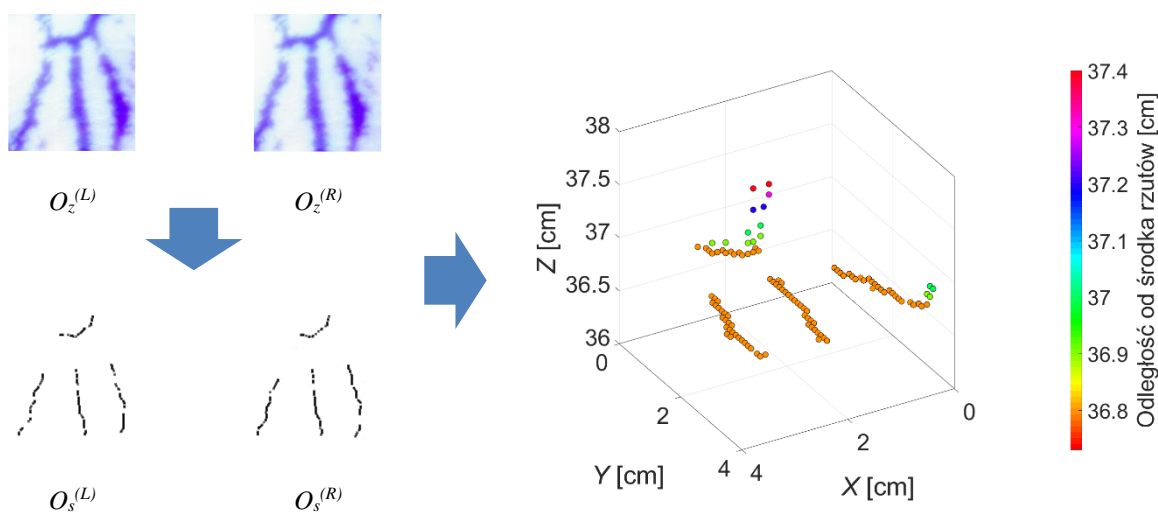
1  for m=1:M do
2    | for i=1:I do
3    | | done=false;
4    | | do
5    | | | LiczDm(m,i);
6    | | | LiczSn;
    
```



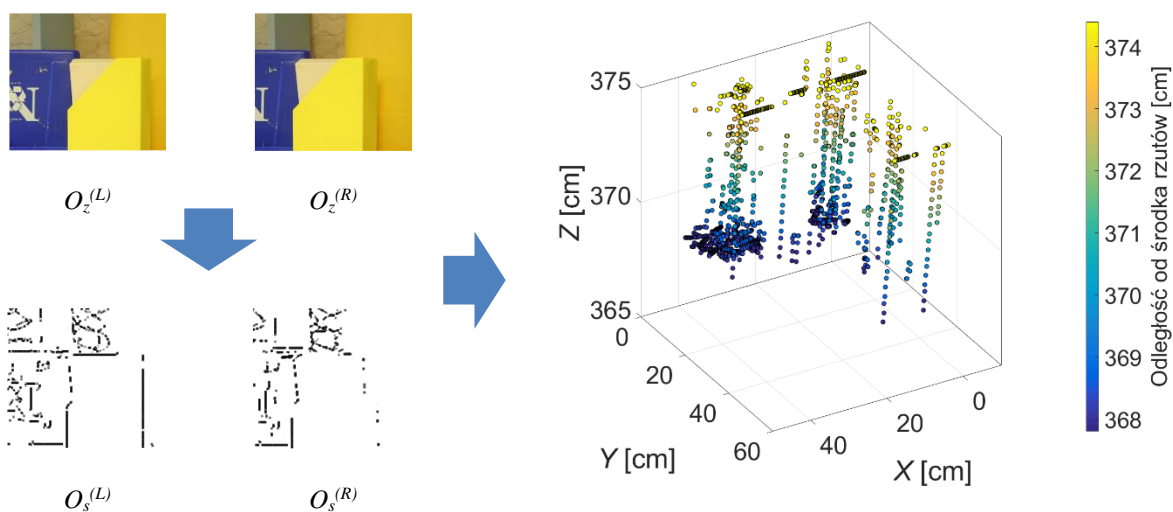
a)



b)



c)



Rys. 6.17. Wynik rekonstrukcji przestrzennej: a) dla obrazu Pir; b) dla obrazu Ves1; c) dla obrazu Plastic.

W macierzy dysparycji pozostaną wartości spełniające nierówność:

$$D(m, n) \leq p_b \cdot D_s \quad (6.39)$$

gdzie  $D_s$  jest medianą macierzy dysparycji.

Wprowadzając parametr  $p_b$  kontrolujemy w jakim stopniu wartość mediany wpływa na proces eliminacji błędnych sparowań.

Po ostatecznym ustaleniu wartości macierzy  $D$ , można przystąpić do obliczenia współrzędnych chmury punktów. Przykładową rekonstrukcję 3D dla wybranych obrazów przedstawiono na rys. 6.17.

W następnym rozdziale przeprowadzona zostanie procedura testowania opracowanych metod poszukiwania odpowiedniości punktów.

## 7. TESTOWANIE OPRACOWANEGO SYSTEMU

Procedurę testową przeprowadzono niezależnie dla każdej metody segmentacji, tj. BGOM, SSSB i BLG testując trzy zaproponowane metody poszukiwania odpowiedniości pomiędzy obrazami, tj. MED, MED-NDD i MED-RGB. Testy przeprowadzono na materiale badawczym przedstawionym w rozdziale 5, tj. obrazach brył stereometrycznych, naczyń podskórnych i obrazów rzeczywistych, co zostało przedstawione na rys. 7.1.

		M A T E R I A Ł   B A D A W C Z Y		
		Bryły stereometryczne	Naczynia podskórne	Middlebury
S E G M E N T A C J A	BGOM	MED	MED	
		MED – NDD	MED – NDD	
		MED – RGB	MED – RGB	
	SSSB	MED	MED	
		MED – NDD	MED – NDD	
		MED – RGB	MED – RGB	
	BLG			MED
				MED – NDD
				MED – RGB

Rys. 7.1 Obszary analizy ilościowej wraz z przyporządkowanymi zastosowanymi metodami segmentacji i parowania obrazów do materiału badawczego.

Zarówno bryły stereometryczne jak i naczynia podskórne zostały wstępnie przetworzone wykonując segmentację BGOM i SSSB. Obrazy rzeczywiste z biblioteki Middlebury wysegmentowano metodą BLG.

Testowanie opracowanego systemu przeprowadzono w oparciu o trzy niezależne kryteria, zdefiniowane w następnym podrozdziale, tj. skuteczność parowania, liczbę błędnych sparowań i czas parowania.

## 7.1. Ilościowa ocena wyników poszukiwania odpowiedniości

W celu oceny działania algorytmu przyjęto trzy kryteria. **Pierwsze kryterium**, zgodnie z [80], dotyczy oceny skuteczności parowania obrazów  $\delta_r$  zdefiniowanym jako:

$$\delta_r = \left( \frac{1}{M \cdot N} \sum_{m=1}^M \sum_{n=1}^N |D(m, n) - D_T(m, n)|^2 \right)^{\frac{1}{2}} \quad (7.1)$$

gdzie

$D$  – obliczona macierz dysparycji,

$D_T$  – wzorcowa macierz dysparycji.

Przyjmując, że korzysta się z dysparycji wzorcowych  $D_T$  przygotowanych dla obrazów *RGB*, odkrywa się ważną cechę powyższej formuły. Jej wynik jest silnie uzależniony od wyników obróbki wstępnej, zwłaszcza segmentacji krawędzi. Im mniej krawędzi zostanie wysegmentowanych, tym więcej jest na obrazie i w macierzy dysparycji obszarów „0”. w skrajnym wypadku skuteczność parowania obrazów będzie zmierzać do średniej różnicy pomiędzy płaszczyzną zerową  $D$ , a wartościami macierzy dysparycji wzorcowej  $D_T$ .

W procesie segmentacji krawędzi naturalnym procesem jest powstawanie obszarów zerowych. Im mniej krawędzi znajduje się na obrazie źródłowym lub im mniej krawędzi zostanie wykrytych, tym więcej obszarów zerowych znajduje się na obrazach wejściowych dla funkcji parowania. Powstała macierz dysparycji per analogia zawiera co najmniej tyle samo obszarów zerowych co lewy obraz wejściowy.

By pozbawić formułę czułości na wynik segmentacji i przystosować ją jedynie do oceny wyników parowania obrazów binarnych, wystarczy dodać warunek, iż formuła działa jedynie dla  $D$  większych od zera, a dysparycji zerowych nie ujmuje się w formule. Tak zmodyfikowana formuła została określona jako  $\delta_p$  i pozwala ocenić jedynie wynik działania funkcji parowania obrazów binarnych.

Jeżeli w procesie binaryzacji uwzględnia się próg, to dla odpowiednio niskiej wartości progu formuła może zwracać niskie wartości  $\delta_p$ , a więc poświadczać wysoką skuteczność metody poszukiwania stereokorespondencji, pomimo braku potwierdzonego wizualnie podobieństwa pomiędzy obrazem binarnym, a źródłowym *RGB*. Formuła może dawać niskie wartości, ponieważ na obrazie, dzięki niskiej wartości progu, znajduje się niewiele elementów.



Dlatego, jeśli nie jesteśmy pewni poprawnego ustalenia progu binaryzacji, można dodatkowo wprowadzić kryterium pozwalające uznać  $\delta_p$  za wartość poprawną. Wartość  $\delta_p$  może być uznana za poprawną tylko dla obrazów z liczbą elementów obrazu  $l_p$ , rozumianych jako piksele o wartości 1, należących do otoczenia  $\varepsilon$  sumy elementów obrazu wzorcowego.

Jako obraz wzorcowy przyjęto wynik operatora Canny'ego dla obrazów źródłowych. Liczba elementów na obrazie krawędziowym powstałym w wyniku działania operatora Canny'ego, w trybie heurystycznym, nazwano liczbą Canny'ego  $l_c$ . Pozwoliło to sformułować wzór:

$$l_c - \varepsilon_c < l_p < l_c + \varepsilon_c. \quad (7.2)$$

Wartość  $\varepsilon_c$  można przyjąć jako najmniejszą wartość otoczenia, dla której występują obrazy binarne o liczbie elementów zbliżonych do liczby Canny'ego. Takie zdefiniowanie  $\varepsilon_c$  może jednak niestety spowodować, iż czasami pula wybranych obrazów będzie zbyt mała, by być reprezentatywną.

Dlatego otoczenie  $\varepsilon_c$  wygodniej jest zdefiniować jako ustaloną z góry część  $l_c$ . w badaniach, tam gdzie zachodziła konieczność wyboru poprawnie zbinaryzowanych obrazów, zastosowano kryterium  $\varepsilon_c$  równego 30 % liczby Canny'ego.

$$0,7 \cdot l_c < l_p < 1,3 \cdot l_c \quad (7.3)$$

W praktyce obrazy binarne, dla których liczba elementów jest zbliżona do liczby Canny'ego, należą do jednego przedziału progu binaryzacji. Wartości  $\delta_p$  pozyskane dla wszystkich obrazów binarnych należących do takiego zakresu progu binaryzacji uznaje się za prawidłowe.

**Drugie kryterium** dotyczy procentowej liczby błędnych sparowań  $\delta_D$  zdefiniowanych zgodnie z [80] jako:

$$\delta_D = \frac{1}{M \cdot N} \sum_{m=1}^M \sum_{n=1}^N L_D^{(L)}(m, n) \quad (7.4)$$

gdzie:

$$L_D^{(L)}(m, n) = \begin{cases} 1 & \text{dla } |D(m, n) - D_T(m, n)| > p_D, \\ 0 & \text{pozostałe} \end{cases} \quad (7.5)$$

$p_D$  – dopuszczalny próg błędu sparowania.

Dopuszczalny próg błędu sparowania  $p_D$  został przyjęty arbitralnie na 2 piksele. Wartość taką ustalono biorąc pod uwagę maksymalną rozdzielczość przestrzenną obrazu oraz rozmiar piksela (dla rozpatrywanego przypadku wielkość piksela nie przekraczała rozmiaru 80  $\mu\text{m}$ ).

Powyższa formuła działa prawidłowo dla obrazów barwnych, gdzie macierz dysparycji ma postać ciągłą. Podobnie jak dla formuły na skuteczność parowania obrazów  $\delta_r$ , również tutaj dostosowano formułę na procentową liczbę błędnych sparowań  $\delta_D$  do oceny parowania obrazów binarnych, podczas korzystania z wzorcowej macierzy dysparycji  $D_T$  w postaci ciągłej, stworzonej do oceny parowania obrazów barwnych lub w odcieniach szarości.

Procentowa liczba błędnych sparowań jest teraz obliczana tylko dla obszarów pokrywających się z wysegmentowanymi krawędziami lewego obrazu, nie zaś dla całej obliczonej macierzy dysparycji  $D$  zawierającej puste obszary reprezentowane przez zera. Tak zdefiniowany błąd określony został jako  $\delta_{Dp}$ .

Gdy nie ma się pewności co do poprawności binaryzacji można skorzystać z formuły (7.2) lub (7.3).

Czas  $t$  przyjęto jako **trzecie kryterium** charakteryzujące algorytm. Czas mierzony jest dokładnie dla operacji działania algorytmu parowania obrazów. Nie wlicza się w czas działania innych operacji, takich jak wczytywanie obrazów, czy ich segmentacja.

Algorytm testowano na komputerze klasy PC z procesorem Intel Core i5 3,1 GHz i 8 GB pamięci RAM. Implementacji dokonano w systemie Matlab.

## 7.2. Wyniki analizy ilościowej

Stosując powyższe kryteria przetestowano wyniki parowania dla zaproponowanych algorytmów parowania MED, MED – NDD i MED – RGB dla obrazów brył stereometrycznych, naczyń podskórnych oraz obrazów rzeczywistych z Middlebury dokonując segmentacji wstępnej obrazów trzema zaproponowanymi technikami, tj. BGOM, SSSB i BLG zgodnie z rys. 7.1. w tabelach pogrubioną czcionką oznaczono wielkości charakterystyczne, tzn. wartość maksymalną i minimalną.

### 7.2.1. Wyniki dla obrazów segmentowanych metodą BGOM

Wyniki dla obrazów własnych zarejestrowanych na potrzeby niniejszej pracy, wysegmentowanych zgodnie z podrozdziałem 6.2.1, przedstawiono w tab. 7.3, tab. 7.4 i tab. 7.5.

Obrazy brył stereometrycznych i naczyń podskórnych dla których wysegmentowano osie, charakteryzują się małą liczbą elementów, co od razu ujawnia się w postaci bardzo krótkich czasów potrzebnych na sparowanie tych obrazów.

Tab. 7.3 Wynik parowania obrazów modeli stereometrycznych i naczyń metodą MED.

	Pir	Cube	Ves1	Ves2	Ves3	Ves4	Ves5	Ves6	Ves7	Ves8	Ves9	Ves10	Średnia
$\delta_p$ [pix]	<b>12,6</b>	8,0	3,0	<b>0,5</b>	3,1	2,1	0,8	8,3	1,7	4,3	4,3	0,7	4,1
$\delta_{Dp}$ [%]	17,5	<b>24,4</b>	7,7	<b>0,0</b>	7,5	11,0	1,6	16,8	6,6	9,2	12,4	0,8	9,6
$t$ [s]	0,02	<b>0,03</b>	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01

Tab. 7.4 Wynik parowania obrazów modeli stereometrycznych i naczyń metodą MED-NDD.

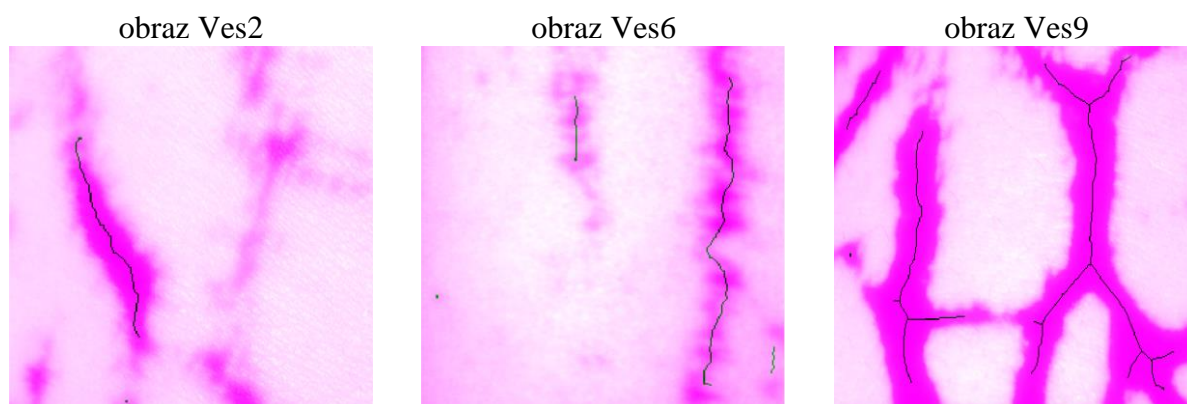
	Pir	Cube	Ves1	Ves2	Ves3	Ves4	Ves5	Ves6	Ves7	Ves8	Ves9	Ves10	średnia
$\delta_p$ [pix]	5,5	5,2	2,6	<b>0,4</b>	2,2	1,1	0,7	<b>8,3</b>	1,1	3,3	3,7	0,7	2,9
$\delta_{Dp}$ [%]	5,3	15,5	6,9	<b>0,0</b>	4,6	4,6	1,7	<b>16,8</b>	5,1	7,4	9,3	0,3	6,5
$t$ [s]	0,03	0,04	0,02	0,02	0,01	0,02	0,01	0,01	<b>0,06</b>	0,02	0,02	0,02	0,02

Tab. 7.5 Wynik parowania obrazów modeli stereometrycznych i naczyń metodą MED-RGB.

	Pir	Cube	Ves1	Ves2	Ves3	Ves4	Ves5	Ves6	Ves7	Ves8	Ves9	Ves10	średnia
$\delta_p$ [pix]	<b>9,3</b>	5,5	1,6	<b>0,5</b>	3,0	2,1	0,7	7,4	1,4	4,1	4,0	0,7	3,4
$\delta_{Dp}$ [%]	10,3	13,9	4,4	<b>0,0</b>	7,5	10,8	1,6	<b>14,4</b>	5,6	5,6	10,6	0,5	7,1
$t$ [s]	0,03	<b>0,07</b>	0,03	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,03

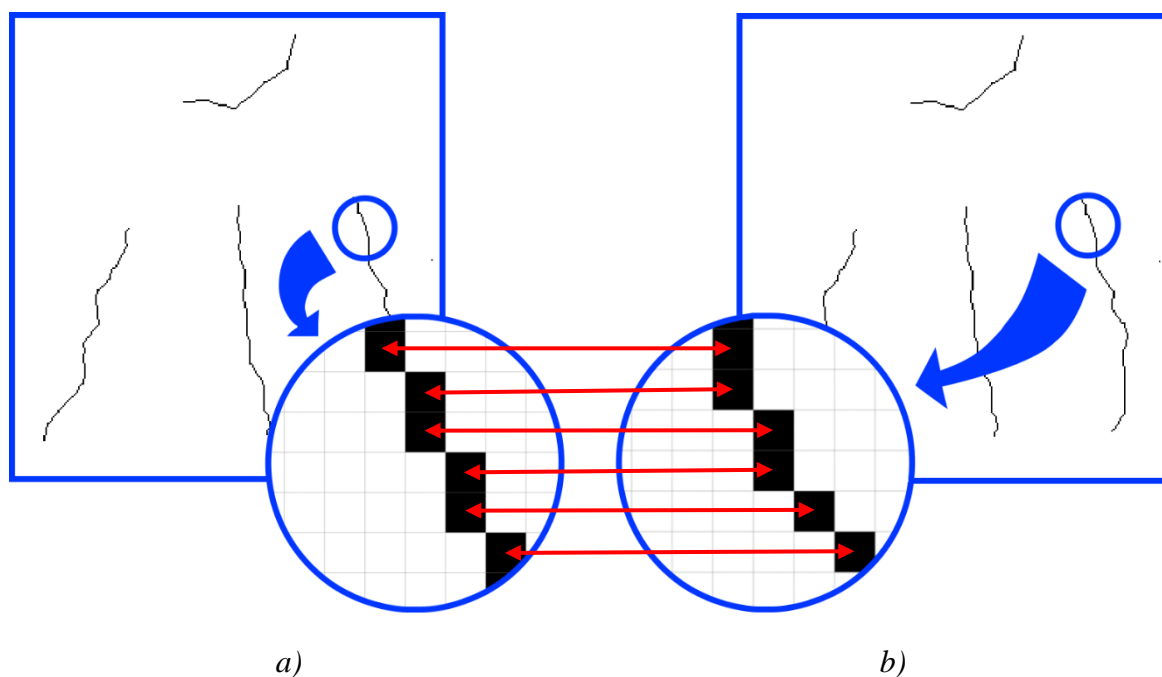
Czasy parowania poniżej 10 ms cechują zarówno metodę najmniej złożoną obliczeniowo MED, jak i jej rozwinięcie MED-NDD. Maksymalny czas 30 ms algorytm MED potrzebował jedynie na sparowanie obrazu Cube. Metoda MED-NDD potrzebowała jedynie na sparowanie obrazu Ves7 czasu 60 ms. Następny bardziej złożony obraz Cube potrzebował 40 ms na sparowanie. Algorytm MED-RGB ogólnie potrzebował mniej niż 20 ms by sparować obrazy. Jedynie parowanie obrazu Cube trwało 70 ms.

Analizując  $\delta_{rp}$  i  $\delta_{Dp}$  należy pamiętać, iż wartość mniejsza oznacza lepszy wynik działania algorytmów parowania.



Rys. 7.2. Nałożenie na siebie obrazów  $O_s^{(L)}$  i  $O_z^{(L)}$ .

Średnia wartość skuteczności parowania obrazów  $\delta_{rp}$  dla metody MED wynosi 4,1 piksela, przy minimalnej wartości 0,5 piksela dla obrazu Ves2 i maksymalnej wartości 12,6 dla obrazu Pir.

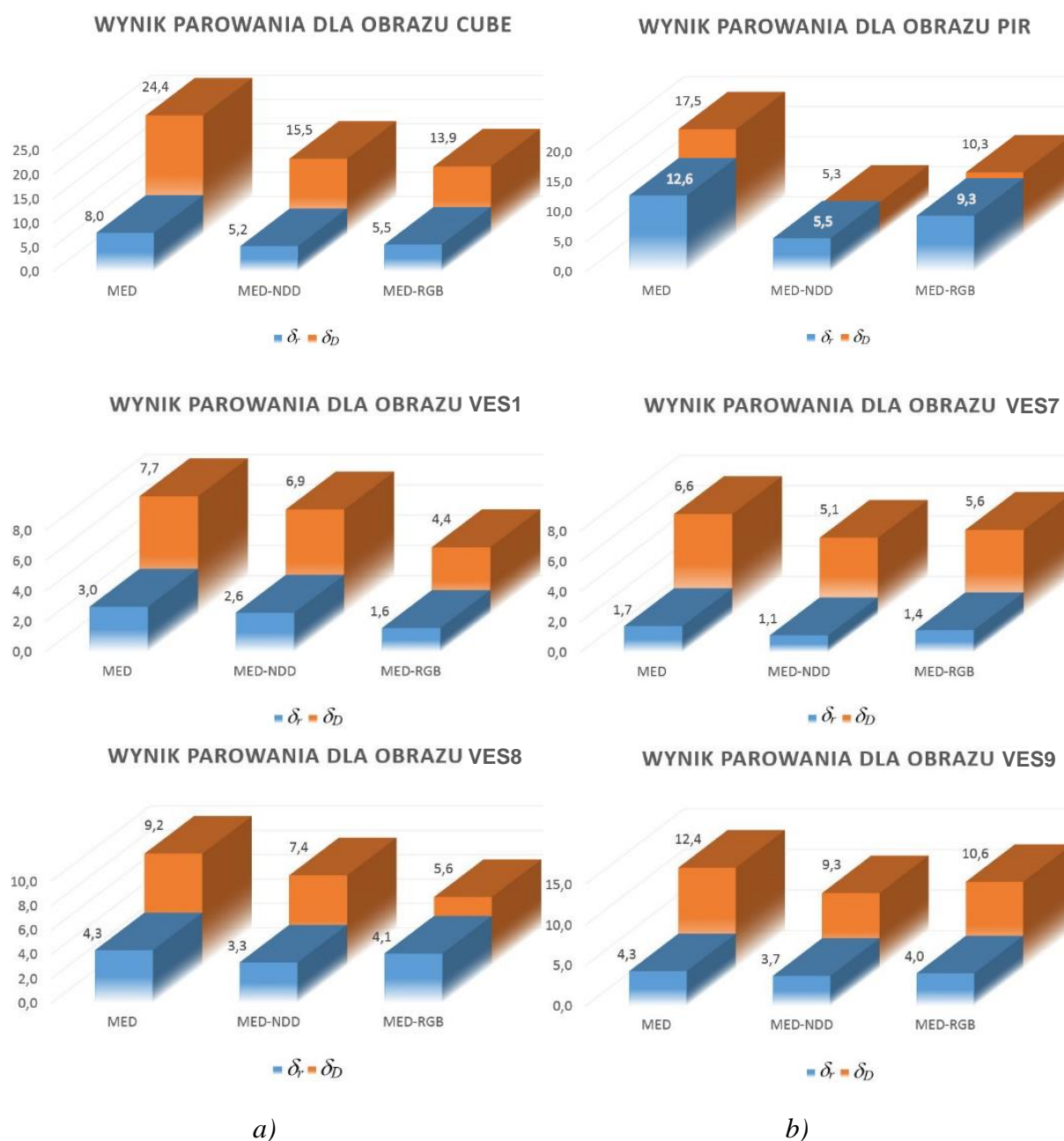


Rys. 7.3. Parowanie punktów tworzących oś naczyń dla zobrazowania Ves1: a) obraz lewy; b) obraz prawy.

Analiza pokazuje iż najlepiej paruje się obraz Ves2. Jest to spowodowane występowaniem na obrazie tylko jednego elementu liniowego zorientowanego wertykalnie, co jest przedstawione na rys. 7.2. Na rys. 7.3 przedstawiony został wynik parowania pomiędzy

lewym i prawym obrazem Ves1 z uwidocznionymi korelacjami pomiędzy poszczególnymi punktami tworzącymi oś naczynia.

Średnia wartość skuteczności parowania obrazów  $\delta_p$  dla metody MED-NDD wynosi 2,9 piksela, przy minimalnej wartości 0,4 piksela dla obrazu Ves2 i maksymalnej wartości 8,3 dla obrazu Ves6.



Rys. 7.4. Porównanie efektów działania algorytmów MED, MED-NDD i MED-RGB dla wybranych zobrażeń modeli stereometrycznych i naczyń podskórnych obrazujących różnice w działaniu algorytmów: w kolumnie a) zebrano obrazy dla których widać wyraźną przewagę algorytmu MED-RGB nad pozostałymi dwoma; w kolumnie b) przedstawiono obrazy na których najlepiej działa algorytm MED-NDD.

Średnia wartość skuteczności parowania obrazów  $\delta_{rp}$  dla metody MED-RGB wynosi 3,4 piksela, przy minimalnej wartości 0,5 piksela dla obrazu Ves2 i maksymalnej wartości 9,3 dla obrazu Pir.

Średnia wartość procentowej liczby błędnych sparowań  $\delta_{dp}$  dla metody MED wynosi 9,6 %. Wartość maksymalna 24,4 % jest wynikiem parowania obrazu Cube, natomiast dla obrazu Ves2 wartość minimalna wyniosła 0 %.

Dla metody MED-NDD średnia wartość  $\delta_{dp}$  wynosi 6,5 %. Parowanie obrazu Ves6 okazało się najtrudniejsze i dało  $\delta_{dp}$  równy 16,8 %, podczas gdy parowanie obrazu Ves2 przebiegło bezbłędnie dając  $\delta_{dp}$  równy 0 %.

Parowanie metodą MED-RGB zakończyło się średnim  $\delta_{dp}$  równym 7,1 %. Maksymalną wartość  $\delta_{dp}$  osiągnął obraz Ves6 z wartością 14,4 %, a wartość minimalną obraz Ves2 z wartością 0 %.

Bardziej złożone obrazy – z większą liczbą elementów, potrzebują więcej czasu na sparowanie. Najszybciej parują się obrazy z małą liczbą obiektów, preferencyjnie o charakterze liniowym, zorientowanych wertykalnie. Analogicznie wartości  $\delta_{rp}$  i  $\delta_{dp}$  osiągają najlepsze wyniki dla najmniej złożonych obrazów. Do takich obrazów zalicza się zobrazenie Ves2, które posiada tylko jeden element liniowy zorientowany wertykalnie.

Na rys. 7.4 widać obraz Ves6, który sprawił najwięcej problemów metodą MED-NDD i MED-RGB. Na tym samym rysunku jest również przedstawiony obraz Ves9, który może się wydawać na pierwszy rzut oka trudniejszy do sparowania niż Ves6. Jednak wyniki dla obrazu Ves9 są znacznie lepsze niż dla Ves6. Jest to spowodowane małą liczbą sąsiednich elementów do porównania dla metody MED-NDD. Obraz Ves9 dysponuje większą liczbą sąsiednich elementów do porównania odległości. Dlatego podjęta decyzja jest bardziej trafna.

Biorąc pod uwagę średnie wartości  $\delta_{rp}$  i  $\delta_{dp}$ , lepszym algorytmem parowania o 15 % od MED-RGB jest MED-NDD. Jednocześnie wyraźnie widać na rys. 7.4 przewagę algorytmów MED-NDD i MED-RGB o 29 % nad algorytmem bazowym dla nich, czyli algorytmem MED. Świadczy to, iż zastosowane metody eliminujące błędne sparowania z macierzy dysparycji, zostały poprawnie sformułowane, a następnie zaimplementowane.

Obrazy naczyń i modeli figur stereometrycznych, po segmentacji przedstawiają małą liczbę elementów, w większości o charakterze liniowym. Im więcej jest na tych obrazach elementów liniowych zorientowanych wertykalnie, tym lepiej sprawują się proponowane metody MED osiągając mniejsze wartości  $\delta_{rp}$  i  $\delta_{dp}$ .

Obrazy naczyń podskórnych należą do zobrazowań o płytkej scenie, którą można uznać za płaską, na których nie dochodzi do przesłaniania się elementów sceny, co tłumaczy dobre wyniki parowania.

### 7.2.2. Wyniki dla obrazów filtrowanych metodą SSSB

Wyniki dla obrazów własnych pozyskanych na potrzeby niniejszej pracy, wysegmentowanych zgodnie z podrozdziałem 6.2.2, przedstawiono w tab. 7.6, tab. 7.7 i tab. 7.8.

Najkrótszy średni czas działania 10 ms cechuje metodę MED, natomiast najdłuższy średni czas 60 ms metodę MED-RGB. Najdłużej parują się obrazy Cube i Ves1, co jest związane z większą liczbą elementów do parowania.

Tab. 7.6 Wynik parowania obrazów brył stereometrycznych i naczyń metodą MED wysegmentowanych wcześniej metodą SSSB.

	Pir	Cube	Ves1	Ves2	Ves3	Ves4	Ves5	Ves6	Ves7	Ves8	Ves9	Ves10	średnia
$\delta_p$ [pix]	<b>15,3</b>	12,7	9,5	9,5	13,5	13,6	8,4	10,7	12,5	10,6	<b>5,8</b>	7,8	10,8
$\delta_{dp}$ [%]	50,7	58,6	57,0	44,9	45,8	<b>66,0</b>	49,7	62,3	62,9	28,5	<b>23,6</b>	53,3	50,3
$t$ [s]	0,02	0,05	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01

Najniższa średnia wartość skuteczności parowania obrazów  $\delta_p$  cechuje algorytm MED-NDD z wartością 6,7 pikseli. Najwyższą wartość średnią 10,8 pikseli wykazał algorytm MED. Należało się spodziewać, iż  $\delta_p$  dla algorytmu MED, będącego bazowym dla dwóch następnych, będzie cechowała się najwyższą wartością. Natomiast algorytmy MED-NDD i MED-RGB cechują się skutecznością na podobnym poziomie.

Tab. 7.7 Wynik parowania obrazów brył stereometrycznych i naczyń metodą MED-NDD wysegmentowanych wcześniej metodą SSSB.

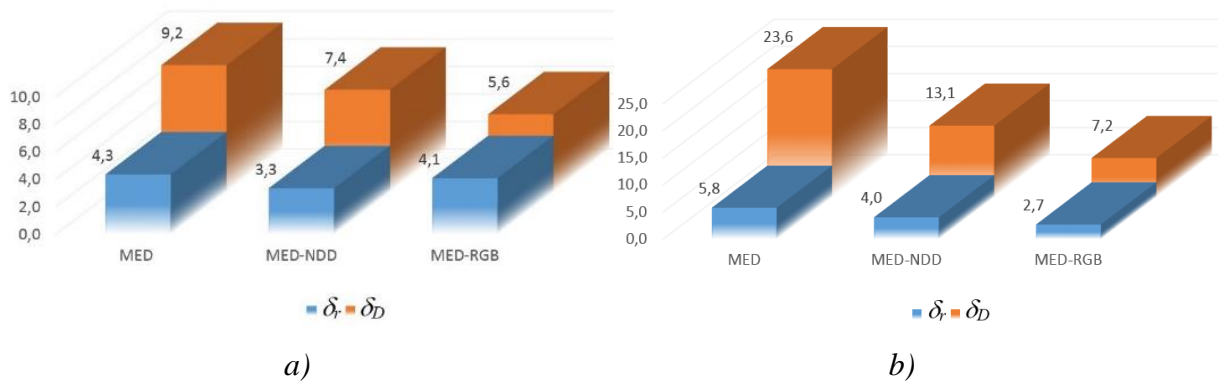
	Pir	Cube	Ves1	Ves2	Ves3	Ves4	Ves5	Ves6	Ves7	Ves8	Ves9	Ves10	średnia
$\delta_p$ [pix]	5,9	7,9	5,4	8,1	<b>10,4</b>	9,1	5,0	7,5	8,2	5,1	<b>4,0</b>	4,3	6,7
$\delta_{dp}$ [%]	<b>8,1</b>	26,5	20,1	27,9	28,2	30,3	22,8	<b>30,2</b>	25,4	11,9	13,1	17,1	21,8
$t$ [s]	0,03	<b>0,08</b>	0,03	0,02	0,02	0,02	0,02	0,02	0,03	0,02	0,02	0,03	0,03

Najniższa średnia wartość procentowej liczby błędnych sparowań  $\delta_D$  cechuje algorytm MED-RGB, z jedynie jedną dziesiątą przewagi nad MED-NDD. Przeszło dwukrotnie wyższą średnią wartość prezentuje algorytm MED.

Tab. 7.8 Wynik parowania obrazów brył stereometrycznych i naczyń metodą MED-RGB wysegmentowanych wcześniej metodą SSSB.

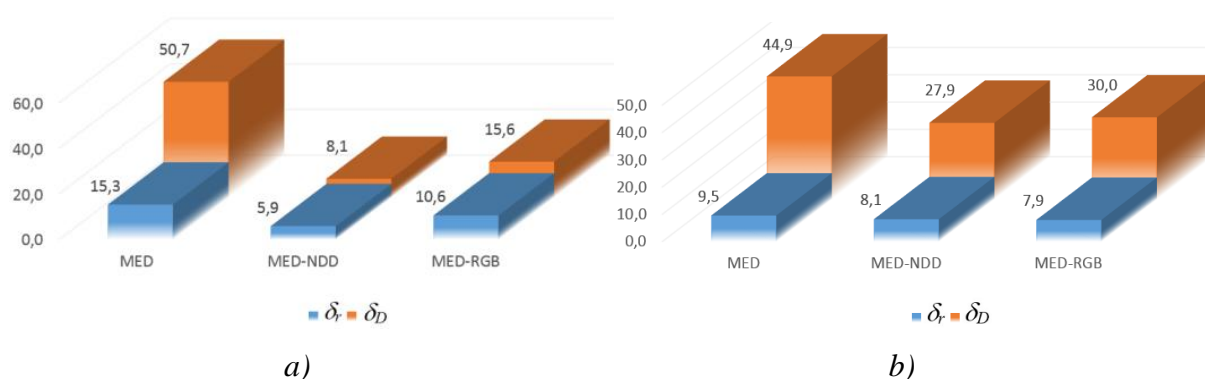
	Pir	Cube	Ves1	Ves2	Ves3	Ves4	Ves5	Ves6	Ves7	Ves8	Ves9	Ves10	średnia
$\delta_T$ [pix]	10,6	8,7	4,4	7,9	14,3	7,5	3,9	6,7	5,6	10,3	2,7	3,0	7,1
$\delta_D$ [%]	15,6	28,4	14,8	30,0	51,3	15,9	13,3	21,8	18,2	31,0	7,2	13,3	21,7
$t$ [s]	0,06	0,2	0,08	0,02	0,06	0,05	0,04	0,03	0,05	0,06	0,04	0,05	0,06

Analizując wzajemną efektywność proponowanych algorytmów, zgodnie z rys. 7.5, rys. 7.6 i rys. 7.7 oczywistym staje się fakt, iż każde zobrazowanie jest na tyle unikalne, że trudno wybrać uniwersalny najlepszy algorytm parowania. Statystycznie rzecz ujmując najlepszą efektywność prezentuje algorytm MED-NDD. Ogólnie metody MED-NDD i MED-RGB dają wyniki  $\delta$  i  $\delta_D$  lepsze o niespełna 50 % od MED.



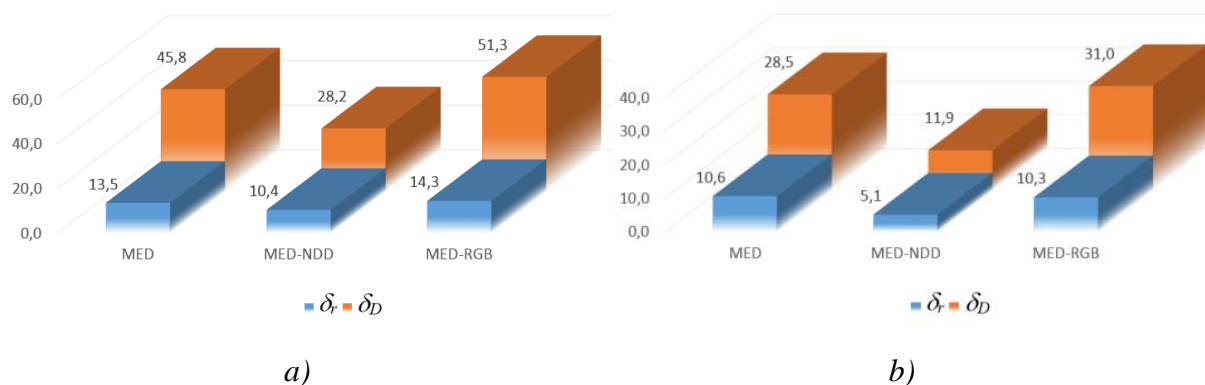
Rys. 7.5. Porównanie efektów działania algorytmów parowania dla dwóch wybranych zobrazowań przedstawiających lepszą efektywność algorytmu MED-RGB: a) wynik parowania dla obrazu Ves8; b) wynik parowania dla obrazu Ves9.





Rys. 7.6. Porównanie efektów działania algorytmów parowania dla dwóch wybranych obrazów przedstawiających lepszą efektywność algorytmu MED-NDD: a) wynik parowania dla obrazu Pir; b) wynik parowania dla obrazu Ves2.

Jednak, każdy z algorytmów MED-NDD i MED-RGB prezentuje najwyższą skuteczność w stosunku do określonych obrazów. Tam gdzie występuje wystarczająco duże zróżnicowanie barwy, najskuteczniejszy jest algorytm MED-RGB. Tam gdzie zróżnicowanie barwy jest nie wystarczające, co jest trudne do oceniania ludzkiemu oku, lepiej prezentuje się MED-NDD.



Rys. 7.7. Dla dwóch obrazów algorytm MED-RGB okazał się mieć gorszą efektywność od algorytmu znacznie prostszego MED: a) wynik parowania dla obrazu Ves3; b) wynik parowania dla obrazu Ves8.

Porównując wyniki z metodą BGOM, metoda SSSB daje o 200 % gorsze wyniki dla  $\delta_r$  i o 300 % dla  $\delta_D$ . Jednak metoda binaryzacji SSSB kryje w sobie duży potencjał polegający na możliwości kontroli procesu binaryzacji przez odpowiedni dobór wartości cech, co zostanie wykorzystane w rozdziale 8.

### 7.2.3. Wyniki dla obrazów filtrowanych metodą BLG

Obrazy testowe z biblioteki Middlebury zostały wysegmentowane metodą BLG zgodnie z opisem w podrozdziale 6.2.3, a wyniki przedstawione w tab. 7.9, tab. 7.10, tab. 7.11. Na rys. 7.8, rys. 7.9 i rys. 7.10 zwizualizowano charakterystyczne przypadki. Wyniki badań dowodzą [74], iż wstępne przetwarzanie tą metodą daje najlepsze wyniki poszukiwania stereokorespondencji dla szerokiej gamy obrazów.

Tab. 7.9 Wynik parowania obrazów Middlebury metodą MED wysegmentowanych metodą BLG.

	Aloe	Baby1	Baby2	Baby3	Bowling1	Bowling2	Cloth1	Cloth2	Cloth3	Cloth4	Flowerpots
$\delta_{rp}$ [pix]	12,9	12,0	<b>11,6</b>	24,8	21,5	28,1	23,7	27,5	30,4	<b>36,7</b>	27,1
$\delta_{Dp}$ [%]	76,1	65,5	75,8	92,4	77,6	81,3	83,0	82,1	<b>99,0</b>	95,6	81,3
$t$ [s]	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,03	0,03	0,02
	Lampshade1	Lampshade2	Mid1	Mid2	Monopoly	Plastic	Rocks1	Rocks2	Wood1	Wood2	średnia
$\delta_{rp}$ [pix]	18,2	18,3	35,8	30,8	19,3	23,2	22,2	20,9	19,2	32,6	22,6
$\delta_{Dp}$ [%]	70,4	73,0	98,0	90,6	86,4	74,5	74,6	71,7	<b>64,1</b>	88,9	77,3
$t$ [s]	0,02	0,02	0,02	0,02	0,02	0,02	0,03	0,02	0,02	0,02	0,02

Algorytm MED prezentuje najniższe czasy działania ze średnim czasem 20 ms. Natomiast średnia wartość  $\delta_{rp}$  jest o 15 % wyższa niż dla metody MED-NDD i o 12 % wyższa niż dla metody MED-RGB. Średnia wartość  $\delta_{Dp}$  jest o 21 % wyższa niż dla algorytmu MED-NDD i o 11 % wyższa niż dla algorytmu MED-RGB. Algorytm MED zgodnie z oczekiwaniami prezentuje gorsze osiągi podczas parowania obrazów od algorytmów MED-NDD i MED-RGB, natomiast najlepszy czas działania, z powodu najmniejszej złożoności obliczeniowej.

Tab. 7.10 Wynik parowania obrazów Middlebury metodą MED-NDD wysegmentowanych metodą BLG.

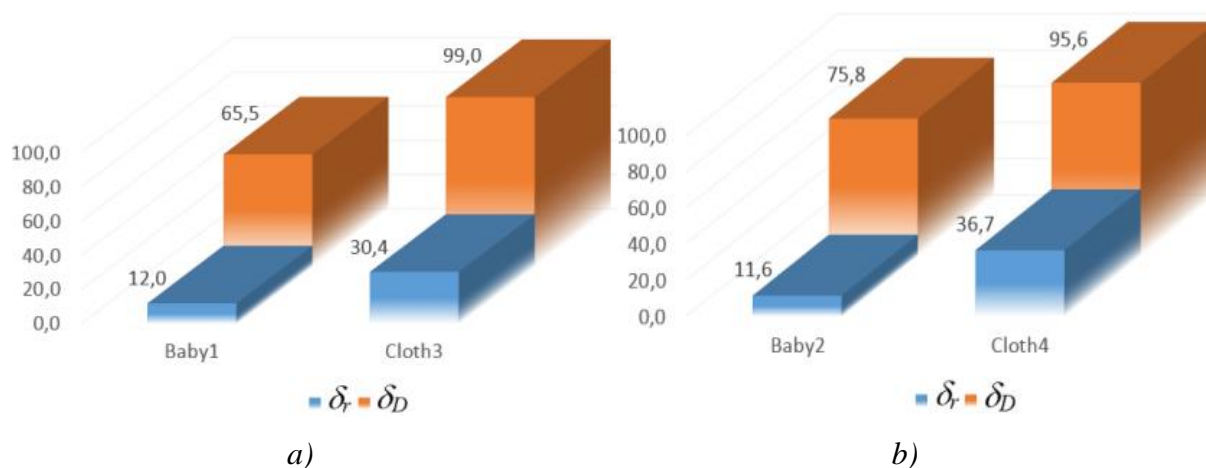
	Aloe	Baby1	Baby2	Baby3	Bowling1	Bowling2	Cloth1	Cloth2	Cloth3	Cloth4	Flowerpots
$\delta_{rp}$ [pix]	11,2	<b>8,9</b>	9,2	21,2	18,8	29,9	16,4	23,3	20,9	25,3	25,6
$\delta_{Dp}$ [%]	55,1	<b>43,3</b>	47,0	67,4	66,7	72,2	51,9	65,8	72,0	68,9	75,2
$t$ [s]	0,23	0,11	0,25	0,36	0,19	0,10	0,18	0,15	<b>0,80</b>	0,56	0,13
	Lampshade1	Lampshade2	Mid1	Mid2	Monopoly	Plastic	Rocks1	Rocks2	Wood1	Wood2	średnia
$\delta_{rp}$ [pix]	16,9	17,9	26,5	25,1	14,9	25,7	19,2	19,2	16,8	<b>30,4</b>	19,2
$\delta_{Dp}$ [%]	59,7	53,6	77,9	77,3	63,9	63,1	63,0	61,7	58,2	<b>81,7</b>	61,1
$t$ [s]	0,10	0,10	0,32	0,19	0,22	0,08	0,13	0,09	<b>0,06</b>	0,07	0,20

Algorytm MED-NDD ze średnią wartością skuteczności parowania obrazów  $\delta_r$  na poziomie 19 pikseli i średnią wartością procentowej liczby błędnych sparowań  $\delta_D$  na poziomie 61 % pretenduje do miana najbardziej skutecznego algorytmu. Optymalizacji wciąż wymaga czas działania algorytmu.

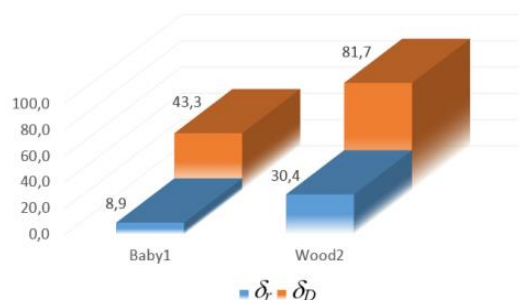
Tab. 7.11 Wynik parowania obrazów Middlebury metodą MED-RGB wysegmentowanych metodą BLG.

	Aloe	Baby1	Baby2	Baby3	Bowling1	Bowling2	Cloth1	Cloth2	Cloth3	Cloth4	Flowerpots
$\delta_r$ [pix]	10,8	<b>9,1</b>	9,7	22,0	19,9	28,3	18,6	24,7	23,0	<b>29,3</b>	25,6
$\delta_D$ [%]	60,1	<b>42,1</b>	60,8	81,0	70,0	78,1	66,7	75,0	79,4	84,2	76,8
$t$ [s]	0,48	0,25	0,64	1,12	0,46	0,18	0,37	0,27	<b>2,43</b>	1,22	0,47
	Lampshade1	Lampshade2	Mid1	Mid2	Monopoly	Plastic	Rocks1	Rocks2	Wood1	Wood2	średnia
$\delta_r$ [pix]	16,8	18,4	26,5	24,6	18,4	26,9	19,0	19,9	18,3	27,9	19,9
$\delta_D$ [%]	67,5	72,6	<b>85,5</b>	80,5	85,3	78,0	68,4	69,3	65,2	77,1	69,1
$t$ [s]	0,28	0,21	1,21	0,67	1,00	0,86	0,43	<b>0,15</b>	0,15	1,44	0,65

Choć w ogólnym ujęciu algorytm MED-RGB plasuje się za algorytmem MED-NDD to dla niektórych obrazów prezentuje niższe wartości skuteczności parowania obrazów  $\delta_r$  (Aloe), a dla innych osiąga bardzo podobne rezultaty, jak MED-NDD (Flowerpots, Lampshade1).

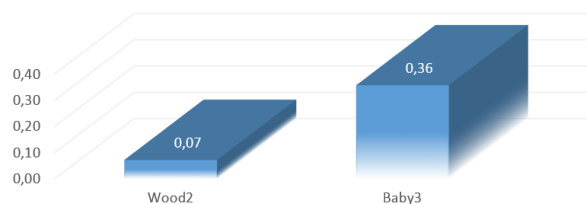


Rys. 7.8. Porównanie wyników dla algorytmu MED.: a) minimalna i maksymalna wartość  $\delta_D$ ; b) minimalna i maksymalna wartość  $\delta_r$ .

MINIMALNA I MAKSYMALNA WARTOŚĆ  $\delta_r$  I  $\delta_D$  DLA METODY MED-NDD

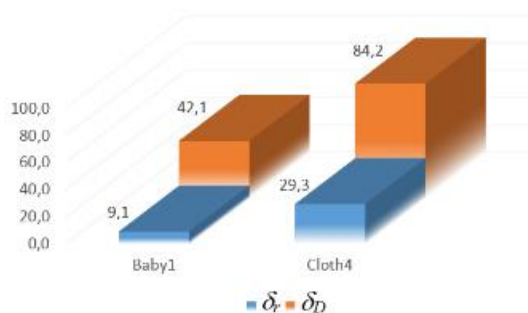
a)

MINIMALNA I MAKSYMALNA WARTOŚĆ CZASU DZIAŁANIA DLA METODY MED-NDD

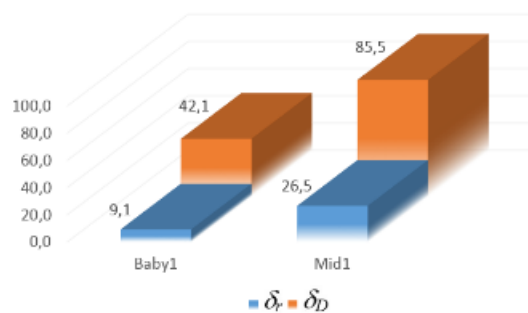


b)

Rys. 7.9. Porównanie wyników dla algorytmu MED-NDD: a) minimalna i maksymalna wartość  $\delta_{rp}$  i  $\delta_{rp}$ ; b) minimalna i maksymalna wartość  $t$ .

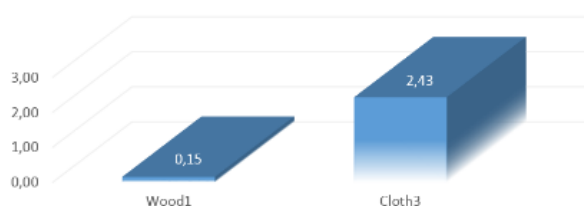
MINIMALNA I MAKSYMALNA WARTOŚĆ  $\delta_r$  DLA METODY MED-RGB

a)

MINIMALNA I MAKSYMALNA WARTOŚĆ  $\delta_D$  DLA METODY MED-RGB

b)

MINIMALNA I MAKSYMALNA WARTOŚĆ CZASU DZIAŁANIA DLA METODY MED-RGB



c)

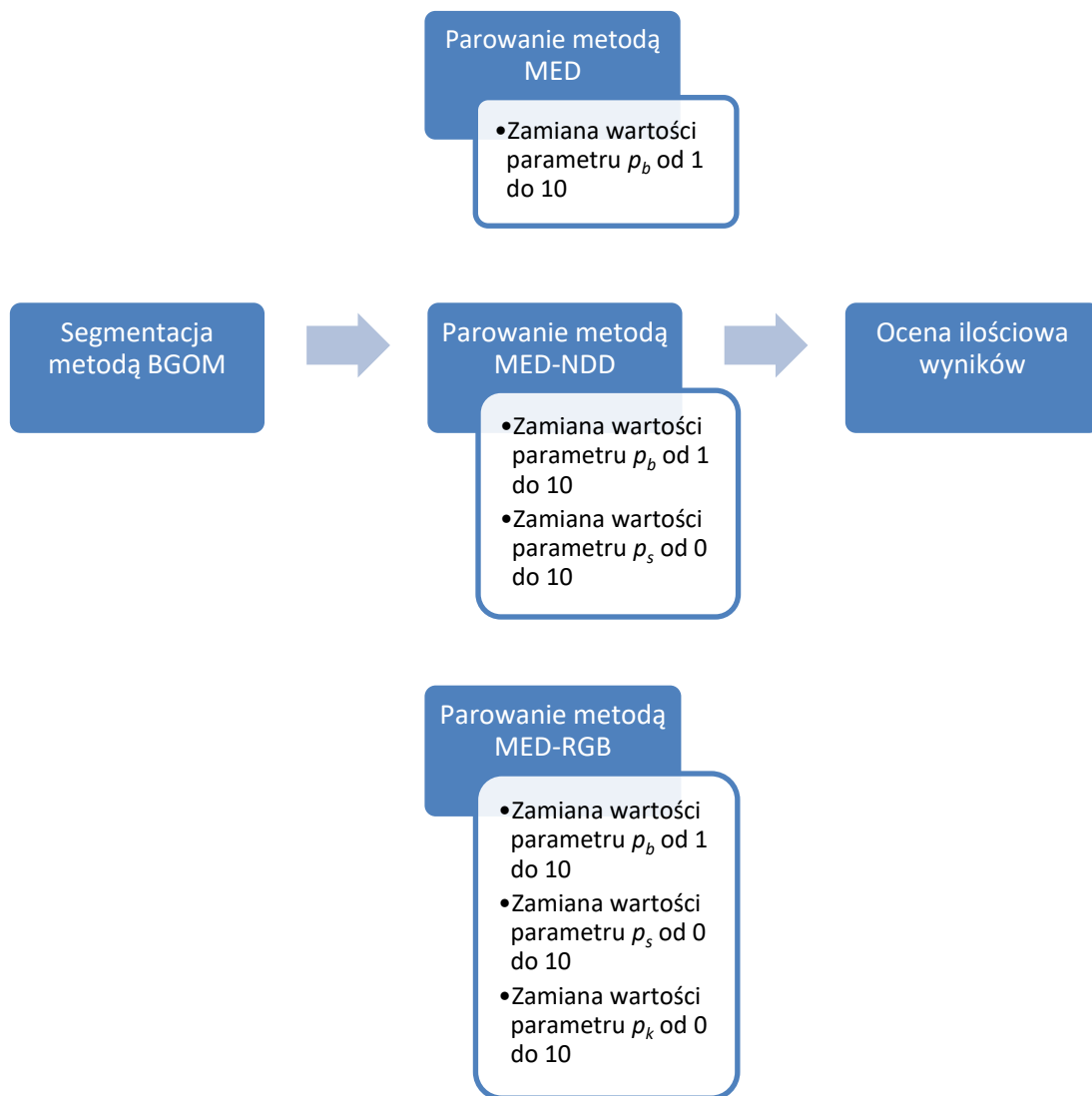
Rys. 7.10. Porównanie wyników dla algorytmu MED-RGB: a) minimalna i maksymalna wartość  $\delta_{rp}$ ; b) minimalna i maksymalna wartość  $\delta_{Dp}$ ; c) minimalna i maksymalna wartość  $t$ .

Reasumując, algorytmy MED, MED-NDD i MED-RGB dla obrazów wysegmentowanych metodą BLG dają o 400 % większe wartości dla  $\delta_r$  i o 500 % większe wartości dla  $\delta_D$ , niż dla obrazów wysegmentowanych metodami BGOM i SSSB dla obrazów naczyń podskórnych. Segmentacja obrazów rzeczywistych innymi metodami niż BLG, nie daje

jednak zadawalających rezultatów. Niższa efektywność parowania dla obrazów rzeczywistych jest zrozumiała, biorąc pod uwagę fakt, iż algorytmy były konstruowane z przeznaczeniem szybkiego parowania obrazów medycznych.

### 7.3. Wrażliwość algorytmu na zmianę parametrów

Wyniki badania wpływu zmiany parametrów na jakość poszukiwania stereokorespondencji przedstawione zostaną na przykładzie dwóch charakterystycznych przypadków, wyraźnie oddających istotę zmian wyników parowania, tj. Aloe oraz Ves1.



Rys. 7.11. Procedura badania wrażliwości algorytmów na zmianę wartości parametrów.

Procedura badawcza przedstawiona została na rys. 7.11. Należy tu zaznaczyć, iż skala i charakter zmiany wyników jest ściśle uzależniony od charakteru zobrażeń, zawartości sceny i naświetlenia.

Wykresy przedstawiają procentową zmianę wartości  $\delta_r$  i  $\delta_D$  w stosunku do pierwszego wyniku. Na osi odciętych przedstawiono wartości parametrów, dla których badano wpływ na jakość stereokorespondencji. Na osi rzędnych przedstawiono procentową zmianę  $\delta_r$  i  $\delta_D$ .

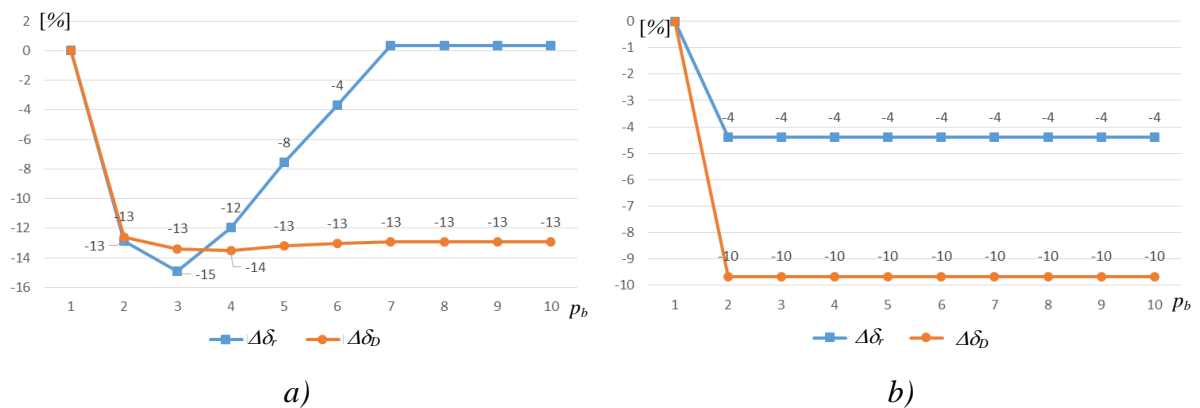
Dla algorytmu MED zbadano parametr odpowiedzialny za filtrację błędnych dysparycji  $p_b$ , który występuje we wszystkich trzech przedstawionych w rozprawie doktorskiej algorytmach parowania. Zbadano zakres wartości parametru od 1 do 10.

Dla algorytmu MED-NDD zbadano parametr  $p_b$  oraz  $p_s$  odpowiadający za decyzję o zaliczeniu odległości do sąsiednich punktów za poprawną bądź błędną. Dla parametru  $p_s$  zbadano zakres wartości od 0 do 10.

Dla algorytmu MED-RGB zbadano parametry  $p_b$ ,  $p_s$  oraz  $p_k$  odpowiadający za zawartość wektora odległości radiometrycznych. Parametr  $p_k$  zbadano dla zakresu wartości od 0 do 10.

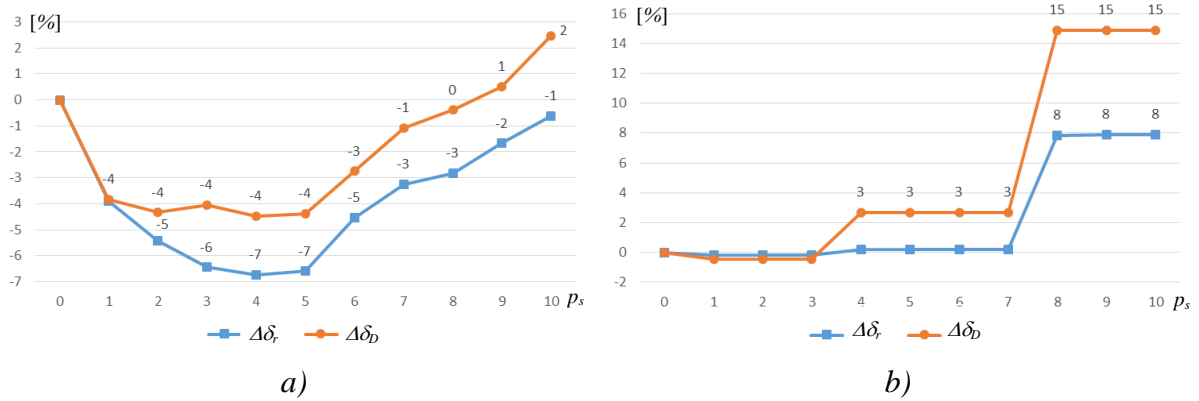
Należy pamiętać, iż im większa jest ujemna wartość procentowa różnicy, tym lepszy jest wynik parowania.

Analizując działanie algorytmu MED dla obrazu Aloe, co jest przedstawiona na rys. 7.12, można zauważyć, iż zarówno wartość  $\delta_r$  jak i  $\delta_D$  na początku maleje wraz ze wzrostem wartości  $p_b$  do wartości 3. Następnie  $\delta_r$  rośnie, by powrócić do wartości początkowej, jaką osiągał dla parametru o wartości 1. Wartość  $\delta_D$ , spada do wartości parametru równego 4, potem nie wiele rośnie i stabilizuje się. Optymalną wartością jest 3. Dla obrazu Ves1 dla wszystkich wartości parametru większych, bądź równych 2, algorytm się stabilizuje, osiągając optymalną dokładność.



Rys. 7.12. Wpływ zmiany parametru  $p_b$  na wartości  $\delta_r$  i  $\delta_D$  dla algorytmu MED: a) dla obrazu Aloe; b) dla obrazu Ves1.

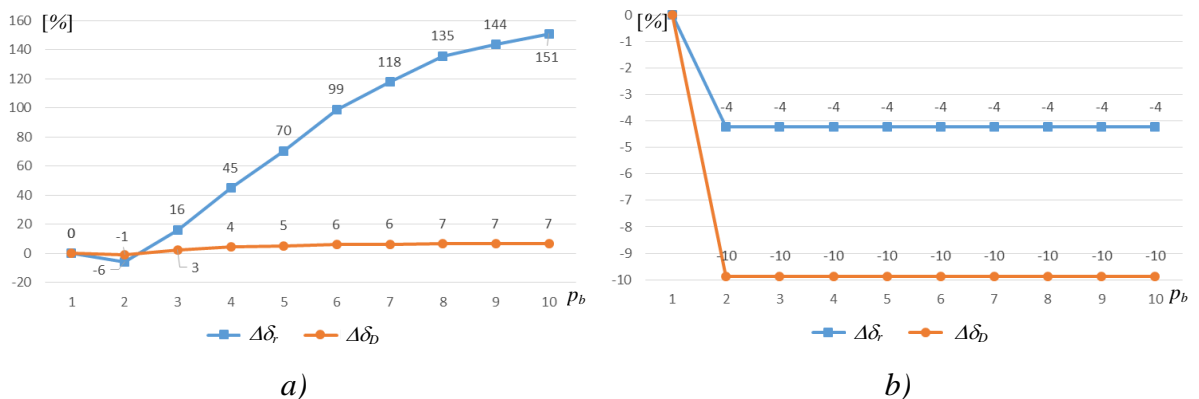
Analizując parametr  $p_s$  występujący w algorytmie MED-NDD dla obrazu Aloe, co zostało przedstawione na rys. 7.13 widać, iż osiąga on optymalny wynik dla wartości 4, by następnie rosnąć.



Rys. 7.13. Wpływ zmiany parametru  $p_s$  na wartości  $\delta_r$  i  $\delta_D$  dla algorytmu MED-NDD: a) dla obrazu Aloe; b) dla obrazu Ves1.

Dla obrazu Ves1 optymalną wartością parametru jest wartość 3. Wartości wyższe powodują schodkowy spadek dokładności parowania obrazów.

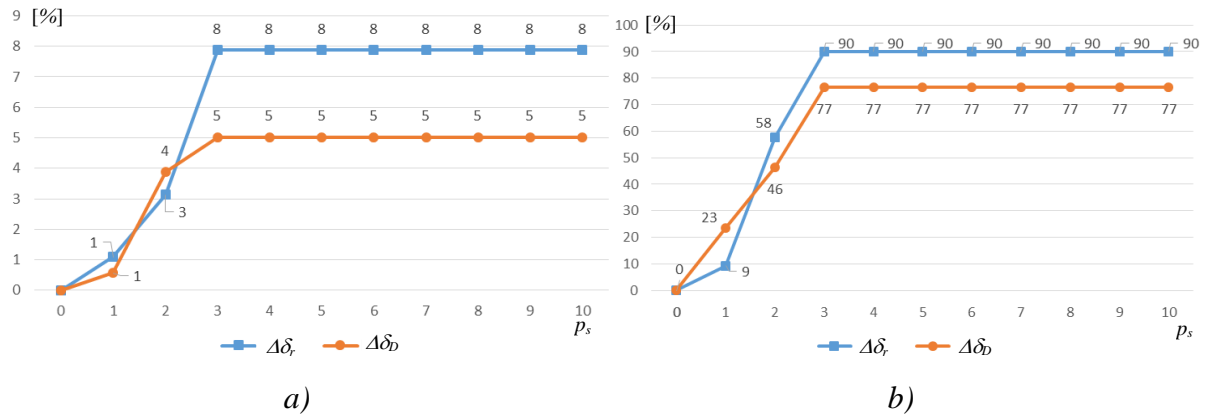
Wynik parowania obrazów dla parametru  $p_b$  dla algorytmu MED-NDD obrazu Aloe, co zostało przedstawione na rys. 7.14, jest optymalny dla wartości parametru 2. Powyżej tej wartości zwiększają się błędy parowania. Dla obrazu Ves1 algorytm osiąga najlepsze rezultaty również dla wartości parametru 2, ale tu inaczej niż wcześniej, wyniki się stabilizują dla wyższych wartości parametru i pozostają nie zmienne przy wzrastającej wartości parametru  $p_b$ .



Rys. 7.14. Wpływ zmiany parametru  $p_b$  na wartości  $\delta_r$  i  $\delta_D$  dla algorytmu MED-NDD: a) dla obrazu Aloe; b) dla obrazu Ves1.

Zmiany parametru  $p_s$  dla algorytmu MED-RGB dają zbliżone rezultaty w charakterze dla obu obrazów testowych Aloe i Ves1, co przedstawione jest na rys. 7.15. Różnica ujawnia się

w skali. Dla obu obrazów wartością optymalną jest 0. Zwiększanie wartości parametru pogarsza wyniki parowania, które stabilizują się dla  $p_s \geq 3$  dla obu obrazów.

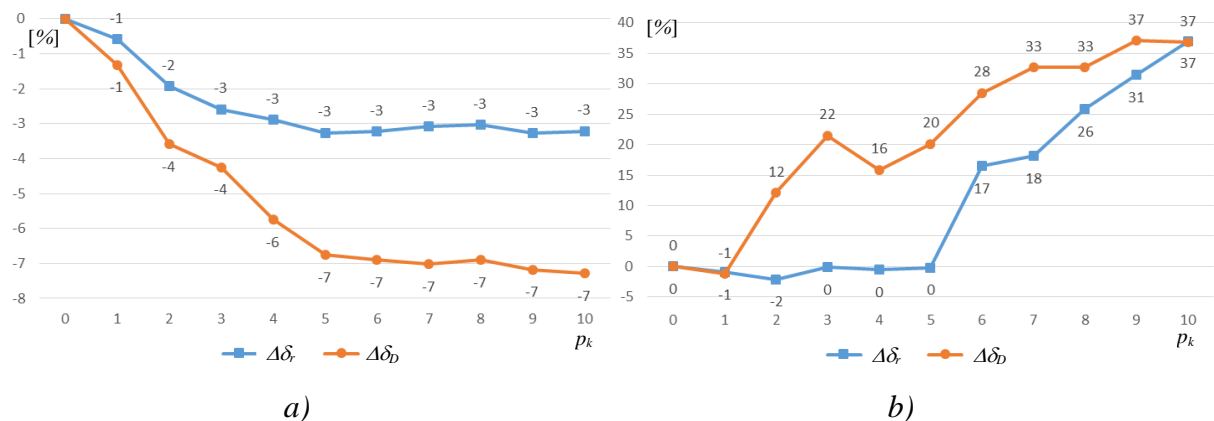


Rys. 7.15. Wpływ zmiany parametru  $p_s$  na wartości  $\delta_r$  i  $\delta_D$  dla algorytmu MED-RGB: a) dla obrazu Aloe; b) dla obrazu Ves1.

Parametr  $p_k$  dla obrazu Aloe wraz ze wzrostem swojej wartości daje lepsze wyniki parowania obrazów. Wynik stabilizuje się dla wartości parametru  $p_k \geq 5$ .

Dla obrazu Ves1 obserwuje się niewielką poprawę wyników parowania dla wartości parametru 1. Dalszy wzrost wartości parametru  $p_k$  powoduje pogorszenie się wyników parowania.

Na rys. 7.17 przedstawiono rekonstrukcję przestrzenną dla wartości minimalnych i maksymalnych parametrów dla wybranych obrazów Pir, Ves1 i Plastic. Na każdym wykresie przedstawiono odpowiednio odmiennymi kolorami wynikowe współrzędne dla przypadku z minimalnymi wartościami parametrów oraz z maksymalnymi wartościami.

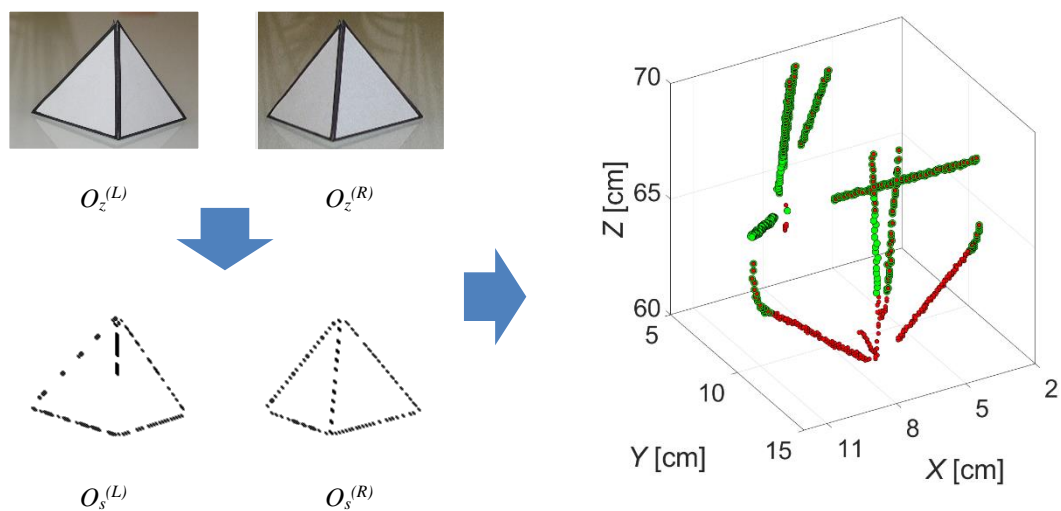


Rys. 7.16. Wpływ zmiany parametru  $p_k$  na wartości  $\delta_r$  i  $\delta_D$  dla algorytmu MED-RGB: a) dla obrazu Aloe; b) dla obrazu Ves1.

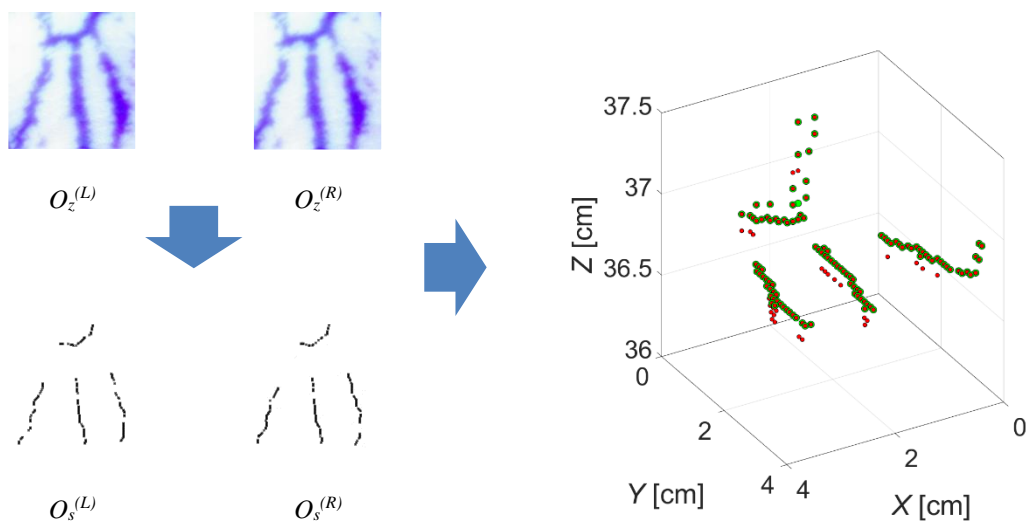
Analizując parametr  $p_b$  dla algorytmu MED-RGB dla obrazu Aloe zgodnie z rys. 7.18, obserwuje się znaczący wzrost wartości  $\Delta\delta_r$  powyżej wartości parametru 2.



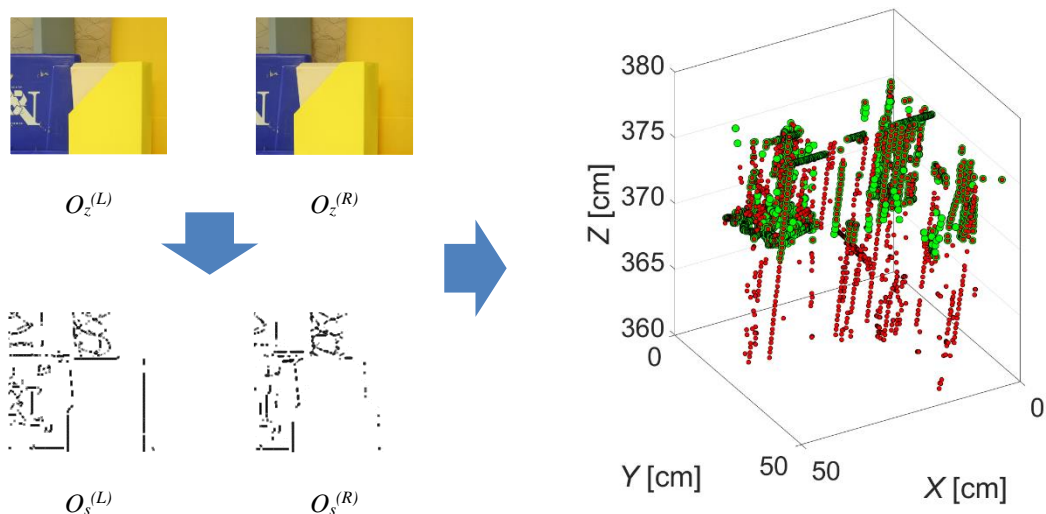
a)



b)

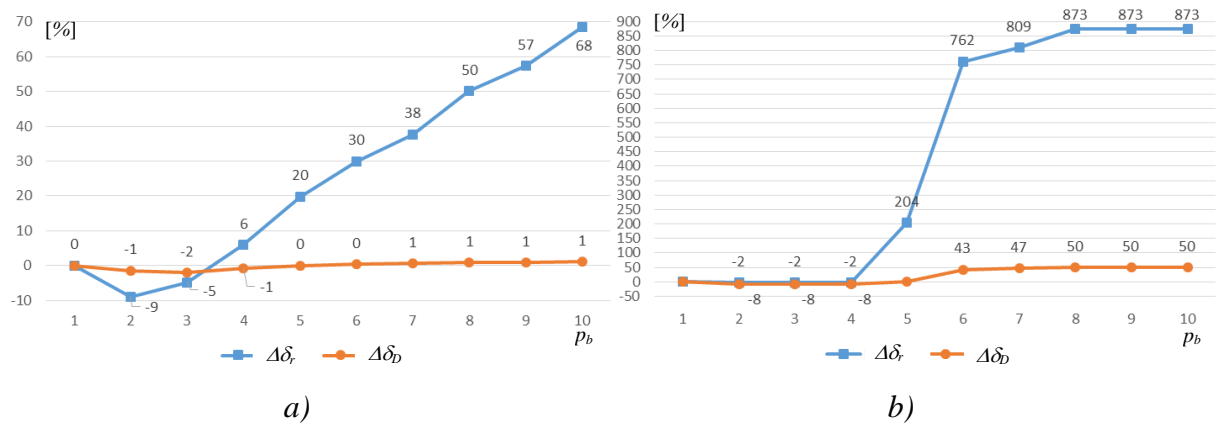


c)



Rys. 7.17. Wynik rekonstrukcji przestrzennej metodą MED-NDD z minimalnymi wartościami parametrów  $p_s = 0$  i  $p_b = 1$  (zielony kolor) oraz maksymalnymi badanymi wartościami  $p_s = 10$  i  $p_b = 10$  (czerwony kolor): a) dla obrazu Pir; b) dla obrazu Ves1; c) dla obrazu Plastic.

Podobny trend obserwuje się dla obrazu Ves1 powyżej wartości 4, jednak w znacznie większej skali.



Rys. 7.18. Wpływ zmiany parametru  $p_b$  na wartości  $\delta_r$  i  $\delta_D$  dla algorytmu MED-RGB: a) dla obrazu Aloe; b) dla obrazu Ves1.

Podsumowując badania wpływu zmiany wartości parametrów algorytmów na wynik parowania obrazów można stwierdzić, iż:

- niewielkim zmianom parametrów mogą towarzyszyć bardzo duże zmiany rezultatów parowania obrazów. Zmiany te mogą sięgać nawet niespełna 900 %.
- statystycznie obrazy o mniejszej liczbie szczegółów są bardziej czułe na zmianę parametrów algorytmów, dając większe różnice w wyniku parowania,
- każdorazowo należy ustalać optymalne wartości parametrów, dopasowując je do poszczególnych zobrażeń.

Ze względu na różnorodność zobrażeń nie da się wybrać konkretnych powszechnie zalecanych wartości parametrów. Można się jednak spodziewać, iż znajdzie się wartość optymalną w zakresie od 0 do 5.

## 8. BADANIA CECH ZNACZĄCYCH

Kontrolowanie cech jest następnym, obok korzystania z wyspecjalizowanych algorytmów parowania, sposobem na polepszenie wyników stereokorespondencji. Dysponowanie gamą wstępnie ustawionych wartości cech, które pozwoliłyby ulepszyć wyniki parowania, byłoby krokiem milowym w rozwoju stereowizji.

Dlatego w niniejszej pracy podjęto badania, by stwierdzić czy kontrola znaczących cech ma wpływ na jakość stereokorespondencji. Oczekiwanym rezultatem miałyby być nie tylko stwierdzenie, czy istnieje taki wpływ, ale i ewentualne ustalenie zestawu uniwersalnych wartości cech, które mogą poprawić proces znajdowania odpowiedniości.

Do testów wpływu cech na jakość poszukiwania odpowiedniości wybrano obrazy rzeczywiste z biblioteki Middlebury. O wyborze zdecydowała różnorodność tych zobrazowań, zarówno pod względem zawartości sceny, jak i warunków radiometrycznych. Dla przestrzeni Hough'a wybrano obrazy brył stereometrycznych. O wyborze zdecydował fakt, iż rzuty tych brył można opisać wyłącznie odcinkami prostymi. Badania przeprowadzono w pięciu przestrzeniach opisanych w podrozdziale 6.3, natomiast wybór badanych cech opisano w podrozdziale 6.4. Cechy znaczące zebrano na rys. 8.1.

Przestrzeń RGB	Przestrzeń HSV	Przestrzeń Lab	Przestrzeń Hough'a	Przestrzeń Canny'ego
<ul style="list-style-type: none"> <li>• <math>p_{RGB}</math></li> <li>• <math>k_R</math></li> <li>• <math>k_G</math></li> <li>• <math>k_B</math></li> </ul>	<ul style="list-style-type: none"> <li>• <math>p_{HSV}</math></li> <li>• <math>k_H</math></li> <li>• <math>k_S</math></li> <li>• <math>k_V</math></li> </ul>	<ul style="list-style-type: none"> <li>• <math>p_{Lab}</math></li> <li>• <math>k_L</math></li> <li>• <math>k_a</math></li> <li>• <math>k_b</math></li> </ul>	<ul style="list-style-type: none"> <li>• <math>l_{max}</math></li> <li>• <math>d_{min}</math></li> <li>• <math>d_{max}</math></li> </ul>	<ul style="list-style-type: none"> <li>• <math>p_{min}</math></li> <li>• <math>p_{max}</math></li> <li>• <math>\sigma</math></li> </ul>

Rys. 8.1. Wykaz badanych cech w poszczególnych przestrzeniach.

Metodę można ogólnie przedstawić w czterech punktach.

1. Ustalenie wartości cech. Zespół cech decydujący o binarnym obrazie podlega ustalonym w iteracjach kontrolowanym zmianom.
2. Stworzenie obrazów binarnych dla stereopary. Formuła zależna jest od przestrzeni w jakiej obraz binarny powstaje. Dla każdego zestawu cech powstają dwa obrazy binarne (stereopara), które są następnie parowane.
3. Poszukiwanie odpowiedniości w stereoparze. w tym celu wykorzystano algorytm MED-NDD (rozdział 6.6). Wywołując algorytm parowania,

wyłączono filtrację błędnych sparowań, by zachować nie zakłóconą postać wyniku.

4. Analiza wyników poszukiwania odpowiedniości. Kontrola jakości parowania obrazów wykonana została przez wyliczenie jednocześnie wartości błędu skuteczności parowania obrazów  $\delta_r$  i  $\delta_{rp}$  oraz wartość błędu procentowej liczby błędnych sparowań  $\delta_D$  i  $\delta_{Dp}$  zgodnie z podrozdziałem 7.1.

Na rys. 8.2 dokonano zobrazowania zaproponowanej metodologii.



Rys. 8.2. Metodologia badania wpływu cech na jakość stereokorespondencji.

W celu zmierzenia wpływu cech na jakość stereokorespondencji wprowadzono dodatkowe kryterium  $\Delta_r$  i  $\Delta_D$ , definiowane jako różnica pomiędzy minimalną i maksymalną wartością  $\delta_r$  i  $\delta_D$  w danej serii pomiarów dla danego obrazu.

$$\Delta_r = \delta_{r \max} - \delta_{r \min} \quad (8.1)$$

$$\Delta_D = \delta_{D \max} - \delta_{D \min} \quad (8.2)$$

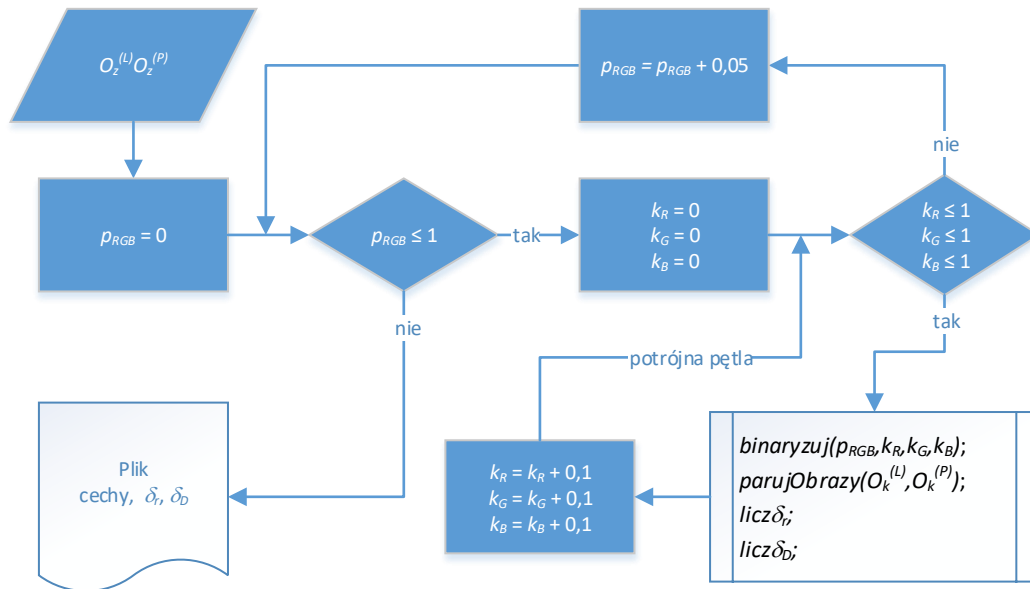
Szczegółowa metodologia dla każdej z przestrzeni oraz wyniki, przedstawione zostaną w kolejnych podrozdziałach.

## 8.1. Analiza cech w przestrzeni RGB

Analiza cech w przestrzeni RGB zostanie przeprowadzona dla 4 cech, tj.  $p_{RGB}$ ,  $k_R$ ,  $k_G$ ,  $k_B$  zgodnie z opisem w rozdziale 6.4. Cechy przyjmują odpowiednio wartości

- od 0 do 1 dla progu  $p_{RGB}$  ze skokiem 0,05,
- od 0 do 1 dla  $k_R$ ,  $k_G$ ,  $k_B$  ze skokiem 0,1.

Operacja binaryzacji, parowania i analizy zamknięta jest w poczwórnej pętli, w której są zmieniane wartości cech decydujących o powstaniu obrazu binarnego, tj.  $p_{RGB}$ ,  $k_R$ ,  $k_G$  i  $k_B$ . Metoda szczegółowo przedstawiona została na schemacie blokowym na rys. 8.3.



Rys. 8.3. Schemat blokowy badania wpływu cech znaczących na parowanie obrazów w przestrzeni RGB. Per analogiam proces zachodzi dla przestrzeni HSV i Lab.

Binaryzacja obrazów przeprowadzona zostanie metodą SSSB opisaną w rozdziale 6.2.2. Jakość znalezionej odpowiedniości punktów zostanie zarejestrowana poprzez wyliczenie dla każdej wartości cechy, wartości błędów  $\delta_r$  i  $\delta_D$  opisanych w rozdziale 7.1.

Dla każdego testowanego obrazu obliczono 27 951 macierzy dysparycji, spośród których wybrano macierz dysparycji, dla której wartości błędów  $\delta_r$ ,  $\delta_p$  i  $\delta_D$ ,  $\delta_{Dp}$  były najniższe i wraz z wartościami cech i nazwą obrazu przedstawiono w tab. 8.12, tab. 8.13, tab. 8.14 i tab. 8.15.

Analogiczna metodologia ma zastosowanie w przypadku przestrzeni HSV i Lab.

Tab. 8.12 Minimalne wartości  $\delta_r$  dla przestrzeni  $RGB$ .

Lp	Nazwa	$\rho$	$k_R$	$k_G$	$k_B$	$\delta_r$	$\delta_D$	$\Delta_r$
<b>1</b>	<b>Aloe</b>	<b>0,95</b>	<b>0,3</b>	<b>0,4</b>	<b>0,7</b>	<b>25</b>	<b>95</b>	<b>6</b>
2	Baby1	0,65	0,0	0,1	1,0	30	100	5
3	Baby2	0,65	0,0	0,1	1,0	30	100	5
4	Baby3	0,95	0,0	0,0	0,7	43	98	3
5	Bowling1	1,0	0,0	0,1-1,0	0,0	44	99	1
6	Bowling2	0,75	0,0	0,1-0,8	0,0	49	100	2
7	Cloth1	0,95	0,2	0,8	0,8	37	97	4
8	Cloth2	0,95	0,4	0,0	0,0	53	98	3
9	Cloth3	0,95	1,0	0,1	0,0	36	97	4
10	Cloth4	0,95	0,8	0,0	0,0	47	98	4
11	Flowerpots	0,95	0,0	0,0	0,4	47	100	4
12	Lampshade1	1,0	0,1-1,0	0,0	0,0	39	99	3
13	Lampshade2	1,0	0,1-1,0	0,0	0,0	41	99	3
14	Midd1	0,95	0,0	0,1	1,0	36	99	5
15	Midd2	0,95	0,0	0,5	0,9	32	99	6
16	Monopoly	0,95	0,0	0,0	0,7	28	98	8
17	Plastic	1,00	0,1-1,0	0,0	0,0	46	99	2
18	Rocks1	0,95	0,0	0,3	0,7	37	97	4
19	Rocks2	0,95	0,1	0,0	0,8	37	98	4
20	Wood1	0,95	0,0	0,0	0,8	41	98	4
21	Wood2	0,95	0,0	0,0	0,7	50	98	2

Minimalne wartości błędów zostały osiągnięte dla obrazu Aloe, osiągając  $\delta_r = 25$  pikseli dla wartości cech  $p_{RGB} = 0,95$ ,  $k_R = 0,3$ ,  $k_G = 0,4$  i  $k_B = 0,7$  i  $\delta_D = 95\%$  dla wartości cech  $p_{RGB} = 0,95$ ,  $k_R = 0,3$ ,  $k_G = 0,4$  i  $k_B = 0,7$ . Wartość progu jest powiązana z intensywnością kanału dla którego występuje najwyższa cecha. Dla wysokich wartości intensywności próg ma również wysoką wartość, osiągając wartość graniczną 1, jak np. dla Bowling1 czy Lampshade1. Dla niższych wartości intensywności również próg ma niższą wartość, jak w przypadku obrazu Baby1 czy Baby2, gdzie osiąga wartość 0,65.

Najczęściej najwyższa wartość cechy towarzyszy kanałowi, dla którego średnia wartość składowej barwy dla całego obrazu w danym kanale jest najniższa. Niższe wartości w danym kanale oznaczają większy kontrast w porównaniu z innymi kanałami, a co za tym idzie większy

wpływ na powstawanie wyraźnych i cienkich krawędzi podczas binaryzacji, które się następnie dobrze parują. Dodatkowo można zauważyć, iż jasność poszczególnych kanałów jest silnie zróżnicowana, tzn. można wizualnie wyróżnić kanał ciemniejszy, jasny i najjaśniejszy. Obrazy robią wrażenie zrównoważonych pod względem liczby obiektów, ich położenia na obrazie i ich barwy. Tezę tą potwierdzają między innymi obrazy Aloe, Baby1, Midd1, Monopoly, Rocks1 i Wood1 (Tab. 8.12).

W badanej próbie obrazów występują również sytuacje dla których pewien zakres wartości cechy (wachlarz) towarzyszy kanałowi, dla którego średnia wartość składowej barwy dla całego obrazu w danym kanale jest najwyższa. Łączy się to z dominacją elementów na obrazie o barwie związanej z kanałem, w którym występuje wachlarz. Obrazy te są zdominowane przez określony obiekt, jego pozycję na obrazie i jego barwę. w efekcie obraz danego kanału jest wypełniony silnym konturem dominujących elementów bez drobnych szczegółów. Wtedy dla każdej wartości cechy z zakresu, binaryzacja daje takie same rezultaty. Przykładami mogą tu być obrazy Bowling2, Lampshade1 lub Plastic (Tab. 8.12).

Jeśli do powyżej opisywanych własności dołączy zbliżona wartość średnia składowej barwy dla dwóch lub trzech kanałów, w jednym z kanałów z najwyższą średnią wartością intensywności pojawia się najwyższa cecha. Może to oznaczać również sytuację, w której dwa kanały zdominowały trzeci. w tym wypadku kanał, który nawet minimalnie dominuje pod względem wysokokontrastowych obiektów, daje lepsze wyniki binaryzacji i co za tym idzie parowania. Przykładem mogą tu być obrazy Cloth1 czy Cloth4 (Tab. 8.12).

Tab. 8.13 Minimalne wartości  $\delta_D$  dla przestrzeni RGB.

Lp	Nazwa	$p$	$k_R$	$k_G$	$k_B$	$\delta_r$	$\delta_D$	$\Delta_D$
<b>1</b>	<b>Aloe</b>	<b>0,95</b>	<b>0,4</b>	<b>0,0</b>	<b>0,0</b>	<b>26</b>	<b>95</b>	<b>5</b>
2	Baby1	0,95	1,0	0,1	0,2	31	96	4
3	Baby2	0,95	1,0	0,1	0,2	31	96	4
4	Baby3	0,95	0,9	0,1	0,2	43	97	3
5	Bowling1	1,0	0,0	0,1-1,0	0,0	44	99	1
6	Bowling2	0,95	0,1	0,0	0,6	50	98	2
7	Cloth1	0,85	0,0	0,0	0,1-0,8	38	97	3
8	Cloth2	0,95	0,1-1,0	0,0	0,0	53	98	2
9	Cloth3	0,95	0,8	0,6	0,4	37	96	4
10	Cloth4	0,90	0,6	0,0	0,2	48	97	3
11	Flowerpots	0,95	0,0	0,7	0,2	48	98	2

12	Lampshade1	0,95	0,0	0,1	0,8	41	98	2
13	Lampshade2	0,95	0,0	0,0	0,7	42	99	1
14	Midd1	1,0	0,0	0,1-1,0	0,0	37	98	2
15	Midd2	1,0	0,0	0,1-1,0	0,0	34	98	2
16	Monopoly	1,00	0,0	0,1-1,0	0,0	32	97	3
17	Plastic	0,95	0,4	0,6	0,0	47	99	1
18	Rocks1	0,95	0,2	0,8	0,0	38	97	3
19	Rocks2	0,90	0,1	0,3	0,8	39	97	3
20	Wood1	0,95	0,1	1,0	0,4	43	98	2
21	Wood2	0,95	0,0	0,5	0,8	50	98	2

Analizując minimalne wartości  $\delta_D$  (Tab. 8.13) można przyjrzeć się trendom intensywności kanałów podobnie, jak podczas analizy wartości  $\delta_r$ . Część obrazów o średnich wartościach intensywności dla poszczególnych kanałów zróżnicowanych względem siebie, robiących wrażenie zrównoważonych zarówno pod względem barwy, liczby i wielkości obiektów, jak i ich rozmieszczenia na obrazie, można zaobserwować większe wartości cech dla składowych barwy o najwyższych wartościach. Przykładem mogą tu być obrazy Baby1, czy Baby2. Tak więc występuje tu odwrotna zależność niż dla  $\delta_r$ .

Tab. 8.14 Minimalne wartości  $\delta_p$  dla przestrzeni RGB.

Lp	Nazwa	$p$	$k_R$	$k_G$	$k_B$	$\delta_r$	$\delta_D$	$\Delta_r$
<b>1</b>	<b>Aloe</b>	<b>0,95</b>	<b>0,3</b>	<b>0,4</b>	<b>0,7</b>	<b>22</b>	<b>82</b>	3
2	Baby1	0,95	0,0	0,1	1,0	33	65	26
3	Baby2	0,95	0,0	0,1	1,0	33	65	28
4	Baby3	0,95	0,0	0,0	0,7	41	88	27
5	Bowling1	1,00	0,0	0,1-1,0	0,0	48	91	15
6	Bowling2	0,95	0,2	0,3	1,0	49	83	10
7	Cloth1	0,95	0,2	0,0	0,4	29	90	3
8	Cloth2	0,95	0,7	0,8	0,1	41	88	2
9	Cloth3	0,95	0,9	0,0	0,1	29	87	13
10	Cloth4	0,95	0,8	0,0	0,0	37	89	2
11	Flowerpots	0,95	0,0	0,0	0,4	41	94	37
12	Lampshade1	1,00	0,0	0,1-1,0	0,0	39	91	23
13	Lampshade2	1,00	0,0	0,1-1,0	0,0	39	94	26
14	Midd1	1,00	0,0	0,1-1,0	0,0	32	90	36



15	Midd2	1,0	0,1-1,0	0,0	0,0	31	86	44
16	Monopoly	0,95	0,0	0,0	0,7	34	73	<b>54</b>
17	Plastic	0,95	0,8	0,4	0,0	28	70	28
18	Rocks1	0,95	0,0	0,4	0,9	32	89	12
19	Rocks2	0,95	0,0	0,3	0,6	31	90	12
20	Wood1	0,95	0,0	0,0	0,8	31	94	36
21	Wood2	0,95	0,3	0,8	0,2	28	81	28

Minimalne wartości błędów  $\delta_{rp}$  i  $\delta_{Dp}$  zostały osiągnięte dla obrazu Aloe, osiągając  $\delta_{rp} = 22$  pikseli dla wartości cech  $p_{RGB} = 0,95$ ,  $k_R = 0,3$ ,  $k_G = 0,4$  i  $k_B = 0,7$  i  $\delta_{Dp} = 58\%$  dla wartości cech  $p_{RGB} = 0,95$ ,  $k_R = 0,0$ ,  $k_G = 0,4$  i  $k_B = 0,2$ .

Dla obrazów dla których:

- średnie wartości intensywności w kanałach nie są już wyraźnie zróżnicowane,
- nie można zaobserwować równowagi pod względem barwy, liczby i wielkości obiektów, jak i ich rozmieszczenia na obrazie,

za rozmieszczenie cech są odpowiedzialne lokalnie dominujące właściwości, trudne do wyłonienia ze zróżnicowanych obrazów.

Tab. 8.15 Minimalne wartości  $\delta_{Dp}$  dla przestrzeni RGB.

Lp	Nazwa	$p$	$k_r$	$k_g$	$k_b$	$\delta_r$	$\delta_b$	$\Delta_D$
1	Aloe	0,90	0,9	0,0	0,0	29	72	10
2	Baby1	0,95	0,9	0,2	0,0	37	59	26
<b>3</b>	<b>Baby2</b>	<b>0,95</b>	<b>0,0</b>	<b>0,4</b>	<b>0,2</b>	<b>36</b>	<b>58</b>	<b>28</b>
4	Baby3	0,95	1,0	0,1	0,0	53	78	14
5	Bowling1	1,0	0,0	0,4	0,7	62	85	7
6	Bowling2	0,95	0,6	0,7	0,0	54	79	11
7	Cloth1	0,85	0,0	0,0	0,1-0,8	31	78	13
8	Cloth2	0,95	0,0	0,8	0,0	42	85	5
9	Cloth3	0,95	0,0	0,8	0,0	36	74	14
10	Cloth4	0,90	0,6	0,0	0,2	37	79	12
11	Flowerpots	0,90	0,0	0,8	0,6	49	80	14
12	Lampshade1	0,95	0,1	0,4	0,9	50	68	25
13	Lampshade2	1,00	0,1	0,7	0,7	59	91	3
14	Midd1	0,90	1,0	0,5	0,5	44	84	6
15	Midd2	0,90	0,5	0,3	1,0	45	83	6

16	Monopoly	0,95	0,6	0,4	0,0	36	68	16
17	Plastic	0,95	0,6	0,2	0,0	29	70	27
18	Rocks1	0,90	0,4	0,8	0,0	38	80	11
19	Rocks2	0,90	0,8	0,1	0,3	41	76	15
20	Wood1	0,90	0,0	0,2	0,5	48	76	18
21	Wood2	0,95	0,2	0,8	0,3	31	81	11

Analiza minimalnych wartości  $\delta_r$  i  $\delta_D$  ogólnie potwierdza wcześniejsze ustalenia dla  $\delta_r$  i  $\delta_D$ . Jednak dla obrazów, gdzie średnie wartości składowych barwy nie są w stosunku do siebie silnie zróżnicowane, wystąpiły przeskoki wartości cech pomiędzy składowymi. Potwierdza to wysoką czułość procesu binaryzacji na odległości pomiędzy średnimi wartościami intensywności poszczególnych składowych barwy *RGB*.

Ogólnie można zaobserwować bardzo wysoką czułość algorytmów parowania na kontrast poszczególnych kanałów, którą wyrażają wartości cech. Znaczenie ma także obszar, na którym dominuje silny kontrast – cały obraz, czy mniejszy obszar, oraz charakter różnic w kontraście, wielkość odstępów pomiędzy wartościami średnimi składowych barwy. w następnym podrozdziale przedstawiona zostanie analiza wpływu cech na jakość odpowiedniości w przestrzeni *HSV*.

## 8.2. Analiza cech w przestrzeni *HSV*

Analiza cech w przestrzeni *HSV* zostanie przeprowadzona dla 4 cech, tj.  $p_{HSV}$ ,  $k_H$ ,  $k_S$ ,  $k_V$  zgodnie z opisem w rozdziale 6.4. Cechy przyjmują odpowiednio wartości

- od 0 do 1 dla progu  $p_{HSV}$  ze skokiem 0,05,
- od 0 do 1 dla  $k_H$ ,  $k_S$ ,  $k_V$  ze skokiem 0,1.

Operacja binaryzacji, parowania i analizy zamknięta jest w poczwórnej pętli, w której są zmieniane wartości cech decydujących o powstaniu obrazu binarnego, tj.  $p_{HSV}$ ,  $k_H$ ,  $k_S$  i  $k_V$ . Metoda szczegółowo przedstawiona została na schemacie blokowym na rys. 8.3.

Binaryzacja obrazów przeprowadzona zostanie metodą SSSB opisaną w rozdziale 6.2.2. Jakość znalezionej odpowiedniości punktów zostanie zarejestrowana poprzez wyliczenie dla każdej wartości cechy wartości błędów  $\delta_r$  i  $\delta_D$  opisanych w rozdziale 7.1.

a)



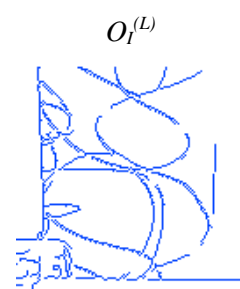
b)



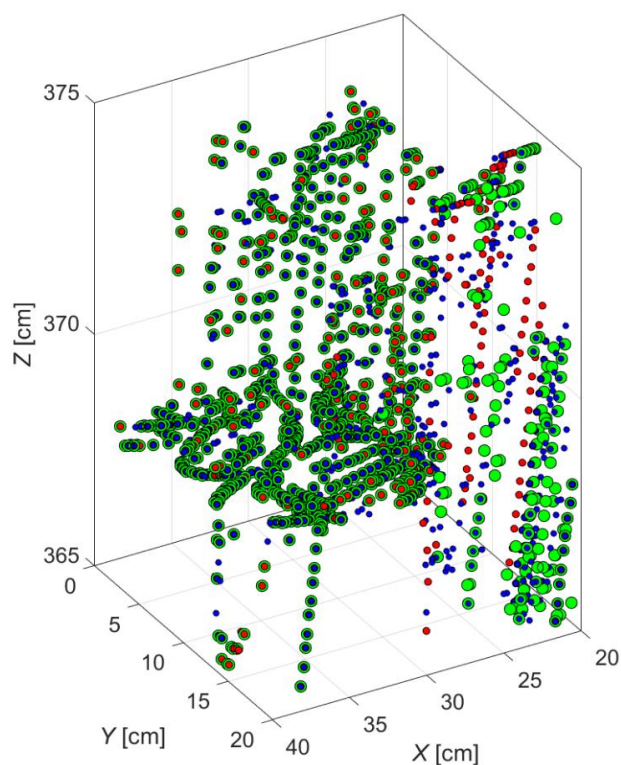
c)



d)



e)



Rys. 8.4. Rekonstrukcja 3D metodą MED-NDD wycinka i obrazu Plastic oznaczonego czerwonym kwadratem, wysegmentowanego metodą SSSB przy zastosowaniu wartości cech minimalizujących błąd  $\delta_r$  (Tab. 8.30): a) obraz wejściowy; b) wysegmentowany obraz z przestrzeni *RGB*; c) wysegmentowany obraz z przestrzeni *HSV*; d) wysegmentowany obraz z przestrzeni *Lab*; e) wynik rekonstrukcji 3D obrazu jednocześnie z trzech przestrzeni barwnych.

Dla każdego testowanego obrazu obliczono 27 951 macierzy dysparycji, spośród których wybrano macierz dysparycji, dla której wartości błędów  $\delta_r$ ,  $\delta_{rp}$  i  $\delta_D$ ,  $\delta_{Dp}$  były najniższe i wraz z wartościami cech i nazwą obrazu przedstawiono w tab. 8.16, tab. 8.17, tab. 8.18 i tab. 8.19.

Tab. 8.16 Minimalne wartości  $\delta_r$  dla przestrzeni *HSV*.

Lp	Nazwa	$p$	$K_H$	$k_S$	$k_V$	$\delta_r$	$\delta_D$	$\Delta_r$
<b>1</b>	<b>Aloe</b>	<b>0,75</b>	<b>0,6</b>	<b>0,0</b>	<b>0,5</b>	<b>26</b>	<b>100</b>	6
2	Baby1	0,70	0,1	0,0	0,3	30	100	5
3	Baby2	0,70	1,0	0,0	0,3	30	100	5
4	Baby3	0,8	0,6	0,4	0,2	43	99	3
5	Bowling1	1,0	0,0	0,0	0,1-0,8	44	99	2
6	Bowling2	0,85	0,8	0,1	0,0	49	100	3
7	Cloth1	0,95	0,2	0,0	0,7	37	97	5
8	Cloth2	0,95	0,4	0,0	0,0	52	98	4
9	Cloth3	0,95	0,0	0,1-1,0	0,0	38	97	2
10	Cloth4	0,95	0,9	0,4	0,0	47	98	4
11	Flowerpots	0,95	0,2	0,9	0,2	49	99	2
12	Lampshade1	1,00	0,0	0,0	0,9	39	99	3
13	Lampshade2	1,00	0,0	0,0	0,7	41	99	4
14	Midd1	0,95	1,0	0,7	0,0	36	99	5
15	Midd2	0,95	0,1	0,5	0,9	32	99	6
16	Monopoly	0,75	0,0	0,9	0,5	28	99	<b>9</b>
17	Plastic	1,00	0,0	0,0	0,1-0,8	46	99	2
18	Rocks1	0,95	0,7	0,0	0,3	38	97	4
19	Rocks2	0,95	0,1	0,0	0,8	38	97	4
20	Wood1	0,95	0,6	0,2	0,9	42	99	4
21	Wood2	0,95	0,4	0,7	0,0	50	99	3

Minimalne wartości błędów zostały osiągnięte dla obrazu Aloe, osiągając  $\delta_r = 26$  pikseli dla wartości cech  $p_{HSV} = 0,75$ ,  $k_H = 0,6$ ,  $k_S = 0,0$  i  $k_V = 0,5$  i  $\delta_D = 95\%$  dla wartości cech  $p_{HSV} = 0,95$ ,  $k_H = 0,8$ ,  $k_S = 0,0$  i  $k_V = 0,3$ . Analizując rozkład cech dla przestrzeni *HSV* zaobserwować można, iż najwyższe wartości cech najczęściej występują dla jasności *V* – sześć razy, pięć razy wyeksponowany jest odcień barwy *H* i tylko trzy razy nasycenie *S*.

Tab. 8.17 Minimalne wartości  $\delta_D$  dla przestrzeni *HSV*.

Lp	Nazwa	$\rho$	$k_H$	$k_S$	$k_V$	$\delta_r$	$\delta_D$	$\Delta_D$
<b>1</b>	<b>Aloe</b>	<b>0,95</b>	<b>0,8</b>	<b>0,0</b>	<b>0,3</b>	<b>28</b>	<b>95</b>	<b>5</b>
2	Baby1	0,95	0,0	0,1	0,9	31	96	4
3	Baby2	0,95	0,0	0,1	0,9	31	96	4
4	Baby3	0,95	0,2	0,0	0,8	44	97	3
5	Bowling1	0,95	0,0	0,1-1,0	0,0	44	98	2
6	Bowling2	0,95	0,3	0,0	1,0	50	98	2
7	Cloth1	0,90	0,1	0,9	0,0	39	96	4
8	Cloth2	0,95	0,8	0,3	0,2	53	98	2
9	Cloth3	0,95	0,0	0,1	0,9	38	97	3
10	Cloth4	0,95	0,9	0,0	0,2	48	97	3
11	Flowerpots	0,95	0,1	1,0	0,0	49	99	1
12	Lampshade1	0,95	0,0	0,1-0,8	0,0	40	99	1
13	Lampshade2	0,95	0,2	0,8	0,0	41	99	1
14	Midd1	1,00	0,0	0,0	0,7	37	98	2
15	Midd2	1,00	0,0	0,0	0,1-0,8	34	98	2
16	Monopoly	1,00	0,0	0,0	0,4	32	97	3
17	Plastic	0,95	0,0	0,0	0,1-0,8	47	99	1
18	Rocks1	0,95	0,7	0,1	0,5	38	97	3
19	Rocks2	0,95	0,6	0,0	0,2	38	97	3
20	Wood1	0,95	0,0	0,1-1,0	0,0	43	98	2
21	Wood2	0,95	0,0	0,1-1,0	0,0	51	99	1

Wartość progu jest wprost proporcjonalna do wyeksponowanej cechy wartości jaskrawości  $V$ . Podczas wyeksponowania wysoką wartością cechy odcienia  $H$  lub nasycenia  $S$  wartość progu jest odpowiednio dostosowana, by podwyższyć kontrast kanału i wydobyć więcej szczegółów.

Tab. 8.18 Minimalne wartości  $\delta_p$  dla przestrzeni *HSV*.

Lp	Nazwa	$\rho$	$k_H$	$k_S$	$k_V$	$\delta_r$	$\delta_D$	$\Delta_r$
<b>1</b>	<b>Aloe</b>	<b>0,95</b>	<b>0,0</b>	<b>0,0</b>	<b>0,1-0,8</b>	<b>24</b>	<b>78</b>	<b>20</b>
2	Baby1	0,95	0,0	0,0	0,9	37	60	8
3	Baby2	0,95	0,8	0,0	0,2	33	60	26
4	Baby3	0,95	0,0	0,1-1,0	0,0	47	84	29
5	Bowling1	0,95	0,0	0,7	0,0	48	83	18

6	Bowling2	0,95	0,0	0,1-1,0	0,0	50	90	24
7	Cloth1	0,95	0,0	0,2	0,8	29	89	7
8	Cloth2	0,95	0,7	0,2	0,1	40	79	5
9	Cloth3	0,95	0,0	0,2	0,8	35	71	14
10	Cloth4	0,95	0,7	0,1	1,0	36	83	9
11	Flowerpots	0,95	0,1-0,8	0,0	0,0	49	89	30
12	Lampshade1	1,00	0,0	0,0	0,9	40	93	24
13	Lampshade2	1,00	0,0	0,0	0,7	39	94	28
14	Midd1	1,00	0,0	0,0	0,8	30	90	36
15	Midd2	1,00	0,0	0,0	0,8	30	86	39
16	Monopoly	0,95	0,0	0,0	0,8	36	67	<b>51</b>
17	Plastic	1,00	0,0	0,0	0,1-0,8	41	97	16
18	Rocks1	0,95	0,7	0,0	0,3	34	87	14
19	Rocks2	0,95	0,1	0,0	0,8	33	85	10
20	Wood1	0,95	0,0	0,1-1,0	0,0	41	82	31
21	Wood2	0,95	0,4	0,7	0,0	29	86	35

Minimalna wartość błędu  $\delta_p$  została osiągnięta dla obrazu Aloe, osiągając  $\delta_p = 24$  pikseli dla wartości cech  $p_{HSV} = 0,75$ ,  $k_H = 0,0$ ,  $k_S = 0,0$  i  $k_V = 0,1-1,0$  i dla obrazu Baby2 osiągając  $\delta_p = 59$  % dla wartości cech  $p_{HSV} = 0,95$ ,  $k_H = 1,0$ ,  $k_S = 0,1$  i  $k_V = 0,5$ . Analizując cechy dla poszczególnych kanałów *HSV*, można wysunąć tezę, iż najwyższe cechy zostały przypisane kanałom o najwyższym kontraście pomiędzy jak najwyższą liczbą szczegółów. Dla niektórych obrazów, kontrast pomiędzy szczegółami obrazu jest wzmacniany przyznając wyższą wartość cechy tylko jednemu kanałowi, najbardziej kontrastowemu, jak np. dla Monopoly (Tab. 8.16) lub Lampshade2 (Tab. 8.17).

Tab. 8.19 Minimalne wartości  $\delta_{Dp}$  dla przestrzeni *HSV*.

Lp	Nazwa	$p$	$k_H$	$k_S$	$k_V$	$\delta_r$	$\delta_D$	$\Delta_D$
1	Aloe	0,95	1,0	0,3	0,5	41	65	17
2	Baby1	0,95	0,1	0,0	1,0	38	60	11
<b>3</b>	<b>Baby2</b>	<b>0,95</b>	<b>1,0</b>	<b>0,1</b>	<b>0,5</b>	<b>38</b>	<b>59</b>	<b>27</b>
4	Baby3	0,95	0,5	0,2	0,6	57	75	20
5	Bowling1	0,95	0,7	0,4	0,0	50	79	15
6	Bowling2	0,95	0,8	0,1	0,6	56	78	17
7	Cloth1	0,95	0,6	0,3	0,1	33	76	17

8	Cloth2	0,90	0,6	0,8	0,1	43	78	13
9	Cloth3	0,95	0,9	0,3	0,8	39	65	19
10	Cloth4	0,95	0,8	0,2	0,2	43	80	11
11	Flowerpots	0,95	0,1-1,0	0,0	0,0	50	89	7
12	Lampshade1	1,00	0,8	0,2	0,0	56	92	2
13	Lampshade2	0,95	0,2	0,6	0,0	40	79	16
14	Midd1	0,95	0,1	0,0	0,8	43	87	7
15	Midd2	0,95	0,7	0,0	1,0	40	86	6
16	Monopoly	0,95	0,7	0,1	0,3	39	66	22
17	Plastic	1,00	1,0	0,1	0,3	54	95	2
18	Rocks1	0,90	0,9	0,0	0,5	42	76	15
19	Rocks2	0,95	1,0	0,1	0,3	41	77	12
20	Wood1	0,95	0,6	0,8	0,1	43	80	15
21	Wood2	0,95	0,9	0,8	0,0	29	86	9

Analiza minimalnych wartości  $\delta_p$  i  $\delta_D$  potwierdza ustalenia dla wcześniejszej analizy dla  $\delta_r$  i  $\delta_D$ . w następnym podrozdziale przedstawiona zostanie analiza wpływu cech na odpowiedniość w przestrzeni *Lab*.

### 8.3. Analiza cech w przestrzeni *Lab*

Analiza cech w przestrzeni *Lab* zostanie przeprowadzona dla 4 cech, tj.  $p_{Lab}$ ,  $k_L$ ,  $k_a$ ,  $k_b$  zgodnie z opisem w rozdziale 6.4. Cechy przyjmują odpowiednio wartości

- od 0 do 1 dla progu  $p_{Lab}$  ze skokiem 0,05,
- od 0 do 1 dla  $k_L$ ,  $k_a$ ,  $k_b$  ze skokiem 0,1.

Operacja binaryzacji, parowania i analizy zamknięta jest w poczwórnej pętli, w której są zmieniane wartości cech decydujących o powstaniu obrazu binarnego, tj.  $p_{Lab}$ ,  $k_L$ ,  $k_a$  i  $k_b$ . Metoda szczegółowo przedstawiona została na schemacie blokowym na rys. 8.3.

Binaryzacja obrazów przeprowadzona zostanie metodą SSSB opisaną w rozdziale 6.2.2. Jakość znalezionej odpowiedniości punktów zostanie zarejestrowana poprzez wyliczenie dla każdej wartości cechy, wartości błędów  $\delta_r$  i  $\delta_D$  opisanych w rozdziale 7.1.

Dla każdego testowanego obrazu obliczono 27 951 macierzy dysparycji, spośród których wybrano macierz dysparycji, dla której wartości błędów  $\delta_r$ ,  $\delta_{rp}$  i  $\delta_D$ ,  $\delta_{Dp}$  były najniższe i wraz z wartościami cech i nazwą obrazu przedstawiono w tab. 8.20, tab. 8.21, tab. 8.22 i tab. 8.23.

Tab. 8.20 Minimalne wartości  $\delta_r$  dla przestrzeni *Lab*.

Lp	Nazwa	$p$	$k_L$	$k_a$	$k_b$	$\delta_r$	$\delta_D$	$\Delta_r$
<b>1</b>	<b>Aloe</b>	<b>0,80</b>	<b>0,6</b>	<b>0,0</b>	<b>0,2</b>	<b>26</b>	<b>100</b>	5
2	Baby1	0,90	0,4	0,6	0,2	30	100	4
3	Baby2	0,90	0,4	0,6	0,2	30	100	4
4	Baby3	0,95	0,0	0,4	0,6	43	99	3
5	Bowling1	0,85	0,6	0,2	0,3	44	100	2
6	Bowling2	0,95	0,2	0,6	0,2	49	100	3
7	Cloth1	0,95	0,1-1,0	0,0	0,0	37	97	4
8	Cloth2	0,95	0,1-1,0	0,0	0,0	54	98	2
9	Cloth3	0,95	0,3	0,8	0,7	38	98	3
10	Cloth4	0,95	0,9	0,1	0,0	48	98	3
11	Flowerpots	0,95	1,0	0,3	0,3	49	99	1
12	Lampshade1	0,95	0,0	0,1	0,1	39	100	3
13	Lampshade2	0,95	0,6	0,1	0,1	41	99	4
14	Midd1	0,90	0,9	0,4	0,1	36	99	5
15	Midd2	0,85	1,0	0,5	0,8	32	99	6
16	Monopoly	0,95	0,0	0,9	0,2	28	99	<b>9</b>
17	Plastic	0,95	1,0	0,0	0,0	47	99	1
18	Rocks1	0,95	0,8	0,1	0,0	37	97	5
19	Rocks2	0,95	1,0	0,0	0,0	38	97	4
20	Wood1	0,95	0,9	0,4	0,2	42	99	4
21	Wood2	0,95	0,1-1,0	0,0	0,0	51	99	2

Minimalne wartości błędów zostały osiągnięte dla obrazu Aloe, osiągając  $\delta_r = 26$  pikseli dla wartości cech  $p_{Lab} = 0,80$ ,  $k_L = 0,6$ ,  $k_a = 0,0$  i  $k_b = 0,2$  i  $\delta_D = 95$  % dla wartości cech  $p_{Lab} = 0,95$ ,  $k_L = 0,1-1,0$ ,  $k_a = 0,0$  i  $k_b = 0,0$ . Analizując rozkład cech w przestrzeni *Lab* dla  $\delta_r$  można jedynie wyciągnąć wniosek, iż najwyższą cechą odznacza się kanał o najwyższej kontrastowości pomiędzy największą liczbą elementów w kadrze.

Nie można znaleźć korelacji pomiędzy wartością progu, a cechami i średnimi wartościami składowych  $L$ ,  $a$ ,  $b$ . Najczęściej występującą wartością optymalną progu jest 0,95.



Nie udało się zaobserwować zależności pomiędzy najwyższymi wartościami cech a średnimi wartościami składowych  $L$ ,  $a$ ,  $b$ .

Tab. 8.21 Minimalne wartości  $\delta_D$  dla przestrzeni  $Lab$ .

Lp	Nazwa	$p$	$k_R$	$k_G$	$k_B$	$\delta_r$	$\delta_D$	$\Delta_D$
1	Aloe	0,95	0,1-1,0	0,0	0,0	26	95	5
2	Baby1	0,95	0,1-1,0	0,0	0,0	31	96	4
3	Baby2	0,95	0,1-1,0	0,0	0,0	31	96	4
4	Baby3	0,95	0,1-1,0	0,0	0,0	44	98	2
5	Bowling1	0,95	0,1-1,0	0,0	0,0	44	99	1
6	Bowling2	0,95	0,1-1,0	0,0	0,0	50	98	2
7	Cloth1	0,95	0,3	0,7	0,0	38	97	3
8	Cloth2	0,95	0,7	0,0	0,2	54	98	2
9	Cloth3	0,95	0,1-1,0	0,0	0,0	38	97	3
10	Cloth4	0,95	0,7	0,5	0,1	48	97	3
11	Flowerpots	0,95	0,1-1,0	0,0	0,0	49	99	1
12	Lampshade1	0,95	0,1-1,0	0,0	0,0	39	99	1
13	Lampshade2	0,95	0,1-1,0	0,0	0,0	42	99	1
14	Midd1	0,95	0,6	0,1	0,0	37	99	1
15	Midd2	0,95	0,8	0,1	0,0	33	99	1
16	Monopoly	0,95	1,0	0,2	0,0	29	97	3
17	Plastic	0,95	1,0	0,0	0,0	47	99	1
18	Rocks1	0,95	1,0	0,2	0,3	38	97	3
19	Rocks2	0,95	0,7	0,1	0,1	38	97	3
20	Wood1	0,95	1,0	0,0	0,0	43	98	2
21	Wood2	0,95	0,1-1,0	0,0	0,0	51	99	1

Na rys. 8.4 przedstawiono rekonstrukcję przestrzenną fragmentu obrazu Plastic dla trzech przestrzeni barwnych dla wartości cech minimalizujących błąd  $\delta_r$ . Chmura punktów dla każdej przestrzeni przedstawiona została odmienną barwą w celu ułatwienia ich wizualnego rozróżnienia.

Analizując rozkład cech dla  $\delta_r$  zauważono dwie tendencje:

- najwyższa wartość cechy charakteryzuje kanały o najwyższej kontrastowości, jak w przypadku Cloth4, Midd1 czy Rocks1,

- wachlarz wartości cech charakteryzuje kanały o najwyższej jasności, jak w przypadku Aloe, Cloth3 czy Lampshade2.

Tab. 8.22 Minimalne wartości  $\delta_{rp}$  dla przestrzeni *Lab*.

Lp	Nazwa	$p$	$k_L$	$k_a$	$k_b$	$\delta_r$	$\delta_D$	$\Delta_r$
<b>1</b>	<b>Aloe</b>	<b>0,95</b>	<b>0,1-1,0</b>	<b>0,0</b>	<b>0,0</b>	<b>25</b>	<b>79</b>	9
2	Baby1	-	-	-	-	-	-	
3	Baby2	0,95	0,1-1,0	0,0	0,0	35	60	4
4	Baby3	0,95	0,8	0,1	0,0	50	78	3
5	Bowling1	1,00	0,4	0,8	0,6	53	92	16
6	Bowling2	0,95	0,6	0,2	0,0	52	78	20
7	Cloth1	0,95	0,8	0,0	0,1	29	91	4
8	Cloth2	0,95	0,7	0,2	0,0	41	82	3
9	Cloth3	0,95	0,1-1,0	0,0	0,0	37	73	4
10	Cloth4	0,95	0,8	0,2	0,3	36	82	2
11	Flowerpots	0,95	0,1-1,0	0,0	0,0	50	82	25
12	Lampshade1	1,00	0,9	0,6	1,0	54	92	14
13	Lampshade2	1,0	0,2	0,9	0,2	53	95	13
14	Midd1	0,95	0,6	0,2	0,1	43	86	27
15	Midd2	0,95	1,0	0,0	0,0	37	86	39
16	Monopoly	0,95	1,0	0,0	0,0	36	68	<b>51</b>
17	Plastic	1,00	0,1	0,9	0,0	50	96	7
18	Rocks1	0,95	0,8	0,1	0,0	34	86	9
19	Rocks2	0,95	1,0	0,0	0,0	33	85	9
20	Wood1	0,95	1,0	0,0	0,0	41	78	32
21	Wood2	1,00	0,3	0,8	0,2	62	94	4

Minimalna wartość błędu  $\delta_{rp}$  została osiągnięta dla obrazu Aloe, osiągając  $\delta_{rp} = 25$  pikseli dla wartości cech  $p_{Lab} = 0,95$ ,  $k_L = 0,1-1,0$ ,  $k_a = 0,0$  i  $k_b = 0,0$  i dla obrazu Baby2 osiągając  $\delta_{rp} = 56$  % dla wartości cech  $p_{Lab} = 0,95$ ,  $k_L = 1,0$ ,  $k_a = 0,3$  i  $k_b = 0,0$ .

Tab. 8.23 Minimalne wartości  $\delta_{Dp}$  dla przestrzeni *Lab*.

Lp	Nazwa	$p$	$k_L$	$k_a$	$k_b$	$\delta_r$	$\delta_D$	$\Delta_D$
1	Aloe	0,95	0,9	0,2	0,8	32	67	12
2	Baby1	-	-	-	-	-	-	
<b>3</b>	<b>Baby2</b>	<b>0,95</b>	<b>1,0</b>	<b>0,3</b>	<b>0,0</b>	<b>38</b>	<b>56</b>	4

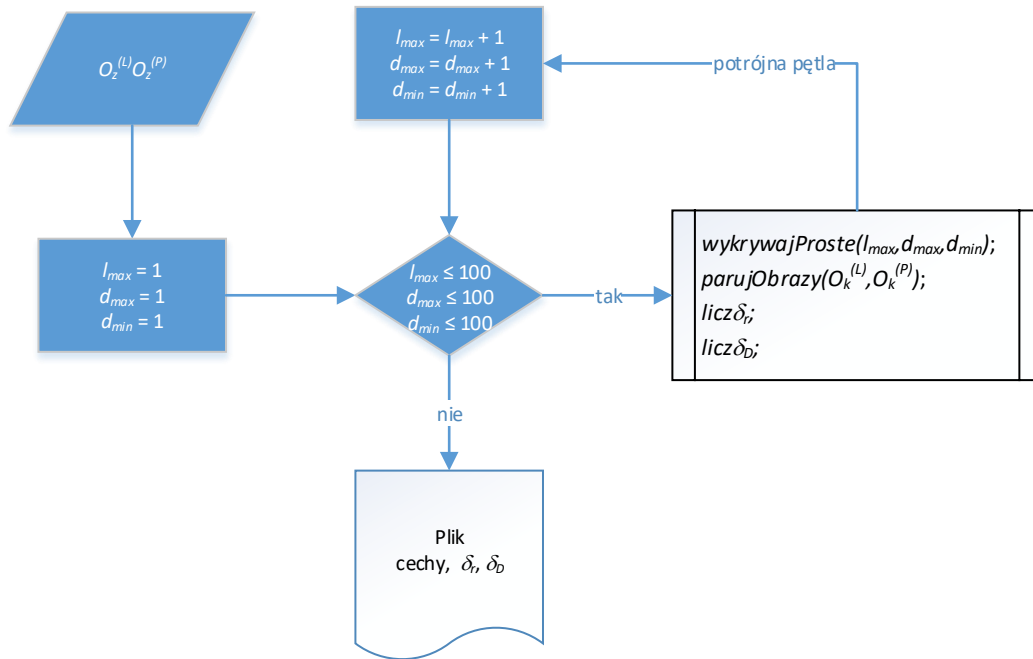
4	Baby3	0,95	0,6	0,2	0,2	52	74	5
5	Bowling1	1,00	0,9	0,3	0,2	58	87	7
6	Bowling2	0,95	0,8	0,2	0,0	54	78	18
7	Cloth1	0,95	0,9	1,0	0,7	31	80	11
8	Cloth2	0,95	0,6	0,0	0,2	43	81	4
9	Cloth3	0,95	0,8	0,3	0,1	40	65	8
10	Cloth4	0,95	0,8	0,7	0,1	38	80	7
11	Flowerpots	0,95	0,1-1,0	0,0	0,0	50	82	14
12	Lampshade1	1,00	1,0	0,1	0,0	58	90	4
13	Lampshade2	1,0	0,9	0,6	0,0	58	91	4
14	Midd1	0,95	1,0	0,8	0,0	46	84	9
15	Midd2	0,95	1,0	0,5	0,0	45	83	9
16	Monopoly	0,95	1,0	0,5	0,0	37	66	<b>22</b>
17	Plastic	1,0	0,8	0,3	0,6	52	95	2
18	Rocks1	0,95	0,7	0,6	0,2	41	74	13
19	Rocks2	0,95	0,9	0,8	0,0	40	75	10
20	Wood1	0,95	1,0	0,0	0,0	41	78	17
21	Wood2	1,00	0,4	0,9	0,3	64	93	3

Analiza cech dla wartości  $\delta_{rp}$  i  $\delta_{dp}$  potwierdza wcześniejsze wnioski wyciągnięte dla wartości  $\delta_r$  i  $\delta_D$ . w następnym podrozdziale przeprowadzona zostanie analiza wpływu cech na jakość odpowiedniości w przestrzeni Hough’a.

#### 8.4. Analiza cech w przestrzeni Hough’a

Analiza cech w przestrzeni Hough’a zostanie przeprowadzona dla 3 cech, tj.  $l_{max}$ ,  $d_{max}$ ,  $d_{min}$  zgodnie z opisem w rozdziale 6.4. Cechy przyjmują odpowiednio wartości od 1 do 100 ze skokiem z inkrementacją o wartość 1.

Badanie cech w przestrzeni Hough’a, wykonano jedynie dla kryteriów dokładności  $\delta_r$  i  $\delta_D$ . Pominięto badanie dla kryteriów  $\delta_{rp}$  i  $\delta_{dp}$  z powodu posiadania własnych dysparycji wzorcowych. Badania przeprowadzono zgodnie ze schematem przedstawionym na rys. 8.5.



Rys. 8.5. Schemat blokowy badania wpływu cech znaczących na parowanie obrazów w przestrzeni Hough'a.

Procedura  $wykrywajProste(l_{max}, d_{max}, d_{min})$  wykonuje transformację Hough'a, a następnie bazując na przestrzeni parametrów rysuje proste odcinki, tworząc obraz binarny. Binaryzacja, parowanie i analiza są zamknięte w potrójnej pętli, w której zmieniają się cechy znaczące dla tej przestrzeni, tj.  $l_{max}$ ,  $d_{max}$  i  $d_{min}$  odpowiedzialne za binaryzację, opisane w podrozdziale 6.4.

Przebieg analizy generuje 1 000 000 macierzy dysparycji, spośród których wybiera się macierz dysparycji dla najniższych wartości kryteriów dokładności. Najniższe wartości kryteriów wraz z wartościami cech i nazwą obrazu przedstawiono w formie tabelarycznej w tab. 8.24, tab. 8.25.

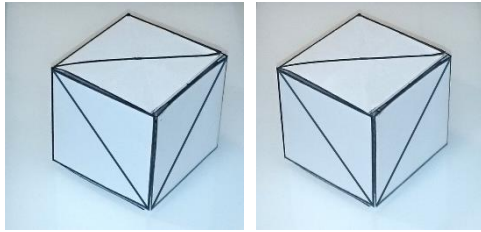
Tab. 8.24 Minimalne wartości  $\delta_r$  dla przestrzeni Hough'a.

Lp	Nazwa	$l_{max}$	$d_{max}$	$d_{min}$	$\delta_r$	$\delta_D$	$\Delta_r$
1	Cube	12-100	29-31	71	14	63	<b>352</b>
2	Pir	<b>5-100</b>	<b>5-100</b>	<b>1-63</b>	<b>9</b>	<b>39</b>	304

Najniższe wartości błędów  $\delta_r$  i  $\delta_D$  osiągnęte są dla cechy  $l_{max}$  (liczby wykrywanych maksimów) od wartości równej liczbie linii na obrazie do maksymalnej badanej liczby 100. Dla przykładu dla Pir osiągnęte są minimalne wartości błędów  $\delta_r$  i  $\delta_D$  od wartości 5, gdyż jest 5 linii

do wykrycia i dla każdej większej liczby, gdyż jest to parametr określający minimalną liczbę maksimów do wykrycia.

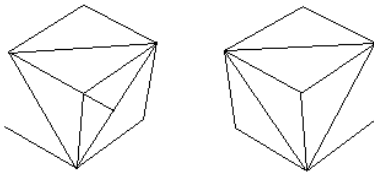
a)



$O_z^{(L)}$

$O_z^{(R)}$

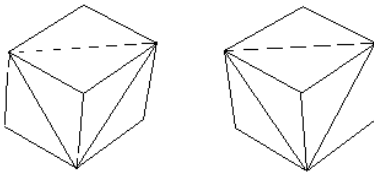
b)



$O^{(L)}$

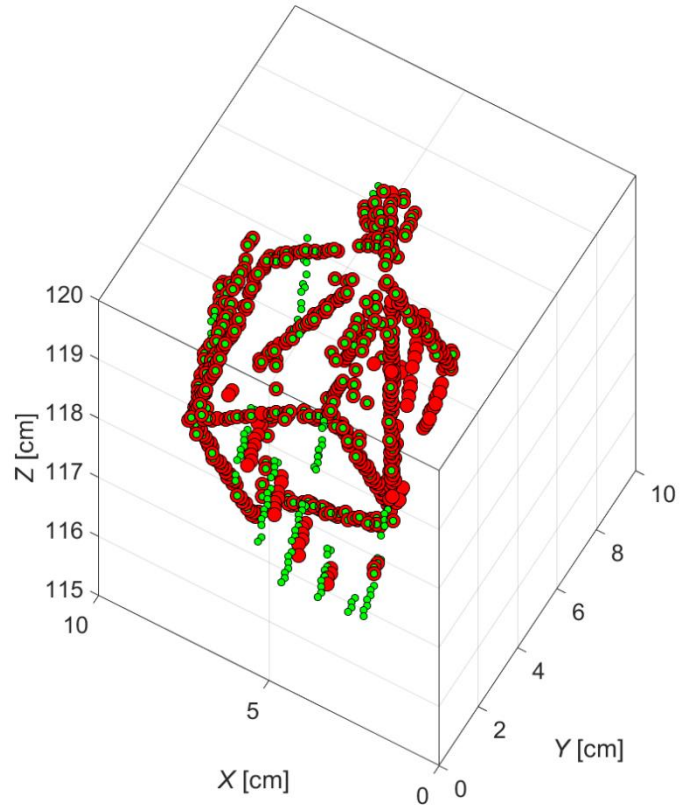
$O^{(R)}$

c)



$O^{(L)}$

$O^{(R)}$



Rys. 8.6. Wynik rekonstrukcji przestrzennej metodą MED-NDD dla przestrzeni Hough'a: a) zrektyfikowane obrazy Cube; b) obrazy dla wartości cech  $l_{max}=12$ ,  $d_{max}=29$  i  $d_{min}=65$  odpowiadających minimalnej wartości  $\delta_r$  (zielony kolor); c) obrazy dla wartości cech  $l_{max}=12$ ,  $d_{max}=5$  i  $d_{min}=1$  odpowiadających minimalnej wartości  $\delta_D$  (czerwony kolor).

Wyniki testów z minimalną długością segmentu linii  $d_{min}$  i maksymalnym odstępem pomiędzy łączonymi liniami  $d_{max}$  świadczą, że najniższe wartości błędów mogą zostać osiągnięte również dla jednopikselowych linii.

Na rys. 8.6 przedstawiono rekonstrukcję przestrzenną odpowiednio dla cech pozwalających osiągnąć minimalną wartość błędu  $\delta_r$  i  $\delta_D$ , barwnie różnicując wyniki.

Podczas analizy obrazu Cube ujawniła się własność, polegająca na lepszym parowaniu linii złożonych z krótszych odcinków, niż długości linii, w tym przypadku linii o długości 8 pikseli.

Tab. 8.25 Minimalne wartości  $\delta_D$  dla przestrzeni Hough'a.

Lp	Nazwa	$l_{max}$	$d_{max}$	$d_{min}$	$\delta_r$	$\delta_D$	$\Delta_D$
1	Cube	12-100	5	1,2,8	19	44	56
2	Pir	5-100	5-100	1-63	9	39	61

Analiza obu kryteriów dokładnościowych, tj.  $\delta_r$  i  $\delta_D$ , pozwala wysunąć podobne wnioski. w następnym podrozdziale przedstawiona zostanie analiza wpływu cech na jakość odpowiedniości dla przestrzeni Canny'ego.

## 8.5. Analiza cech w przestrzeni Canny'ego

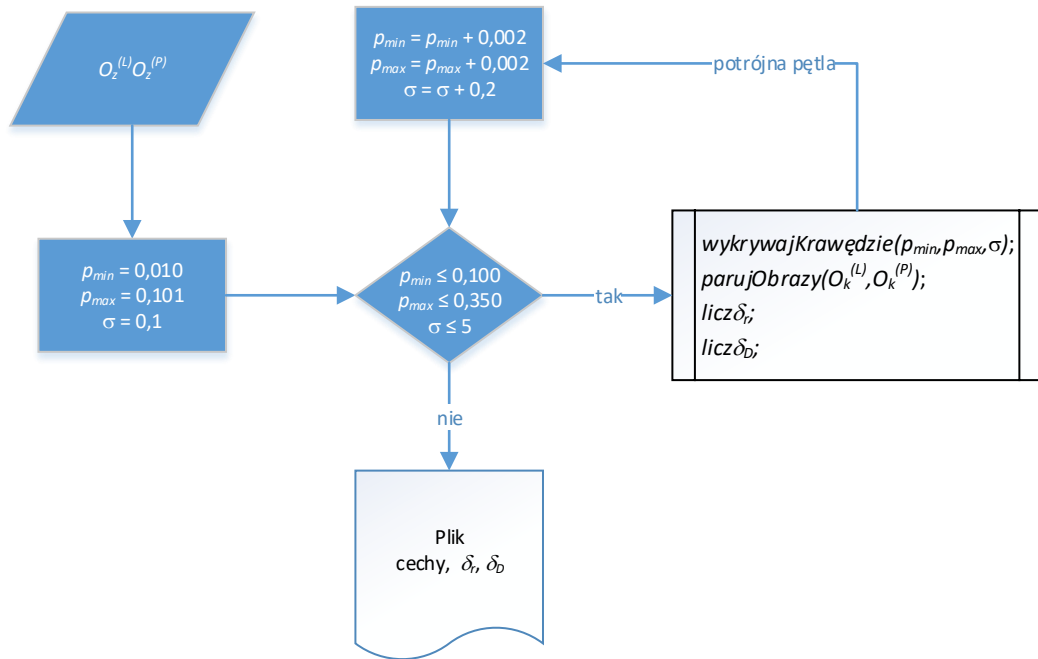
Analiza cech w przestrzeni Canny'ego zostanie przeprowadzona dla 3 cech, tj.  $p_{min}$ ,  $p_{max}$ ,  $\sigma$  zgodnie z opisem w rozdziale 6.4. Cechy przyjmują odpowiednio wartości

- $p_{min}$  od 0,010 do 0,100 ze skokiem z inkrementacją o wartość 0,002,
- $p_{max}$  od 0,101 do 0,350 ze skokiem z inkrementacją o wartość 0,002,
- $\sigma$  od 0,1 do 5 ze skokiem z inkrementacją o wartość 0,2.

Podczas badania wpływu cech na jakość odpowiedniości w przestrzeni Canny'ego, zarejestrowano wartości  $\delta_r$  i  $\delta_D$  oraz  $\delta_{rp}$  i  $\delta_{Dp}$  dla każdego obrazu testowego, zgodnie ze schematem przedstawionym na rys. 8.7. Procedura *wykrywajKrawędzie*( $p_{min}, p_{max}, \sigma$ ) binaryzuje obraz operatorem Canny'ego. Proces binaryzacji, parowania i analizy jest zamknięty w potrójnej pętli, w której iterowane są cechy odpowiedzialne za powstanie obrazu binarnego, tj.  $p_{min}$ ,  $p_{max}$  i  $\sigma$  opisane w podrozdziale 6.4.

Każdy przebieg wygenerował 143 750 macierzy dysparycji, spośród których wybrano macierze dysparycji dla minimalnych wartości kryteriów i wraz z wartościami cech i nazwą obrazu przedstawiono w formie tabelarycznej w tab. 8.26, tab. 8.27, tab. 8.28 i tab. 8.29.

Minimalna wartość  $\delta_r$  uzyskano dla obrazu Aloe, która wynosi 25 pikseli dla wartości cech  $p_{min}=0,036$ ,  $p_{max}=0,101$ ,  $\sigma=0,1$  (Tab. 8.26). Minimalna wartość  $\delta_{rp}$  uzyskano dla obrazu Aloe, która wynosi 22 piksele dla wartości cech  $p_{min}=0,098$ ,  $p_{max}=0,259$ ,  $\sigma=0,5$  (Tab. 8.28).



Rys. 8.7. Schemat blokowy badania wpływu cech znaczących na parowanie obrazów w przestrzeni operatora Canny’ego.

Minimalną wartość  $\delta_D$  uzyskano dla obrazu Baby1, która wynosi 95 procent dla wartości cech  $p_{min}=0,064$ ,  $p_{max}=0,101$ ,  $\sigma=1,1$  (Tab. 8.26). Minimalną wartość  $\delta_{Dp}$  uzyskano dla obrazu Baby1, która wynosi 63 procent dla wartości cech  $p_{min}=0,090$ ,  $p_{max}=0,109$ ,  $\sigma=1,7$  (Tab. 8.28).

Tab. 8.26 Minimalne wartości  $\delta_r$  dla operatora Canny’ego.

Lp	Nazwa	$p_{min}$	$p_{max}$	$\sigma$	$\delta_r$	$\delta_D$	$\Delta_r$
<b>1</b>	<b>Aloe</b>	<b>0,036</b>	<b>0,101</b>	<b>0,1</b>	<b>25</b>	<b>97</b>	2
2	Baby1	0,040	0,101	0,9	29	96	4
3	Baby2	0,016	0,105	0,3	31	96	7
4	Baby3	0,010	0,103	0,3	42	98	3
5	Bowling1	0,092	0,345	2,1	44	100	1
6	Bowling2	0,016	0,101	0,3	49	98	1
7	Cloth1	0,035	0,231	0,5	37	99	2
8	Cloth2	0,010	0,101	0,5	52	99	<b>8</b>
9	Cloth3	0,010	0,101	0,1	36	99	2
10	Cloth4	0,010	0,139	0,7	47	99	4
11	Flowerpots	0,010	0,103	0,5	48	99	2
12	Lampshade1	0,088	0,197	2,3	39	99	6

13	Lampshade2	0,046	0,149	2,7	41	99	6
14	Midd1	0,014	0,105	0,5	36	99	1
15	Midd2	0,012	0,101	0,5	32	99	2
16	Monopoly	0,010	0,101	0,3	28	98	2
17	Plastic	0,012	0,121	0,1	47	99	0
18	Rocks1	0,012	0,101	0,3	37	98	2
19	Rocks2	0,020	0,103	0,1	37	99	3
20	Wood1	0,014	0,103	0,5	41	98	5
21	Wood2	0,012	0,101	0,1	50	99	5

Analizując cechy dla minimalnych wartości  $\delta_r$  formułuje się następujące wnioski:

- dla obrazowań zawierających drobne elementy wartości  $p_{min}$  i  $\sigma$  utrzymują się na niskim poziomie, bliskim wartości wyjściowej pętli obliczeniowej,
- dla obrazowań, na których występują dominujące wielkością jednobarwne elementy  $p_{min}$  i  $\sigma$  utrzymują się na znacznie wyższym poziomie, co najmniej cztery razy większym, niż dla drobnych elementów.

Tab. 8.27 Minimalne wartości  $\delta_D$  dla operatora Canny'ego.

Lp	Nazwa	$p_{min}$	$p_{max}$	$\sigma$	$\delta_r$	$\delta_D$	$\Delta_D$
1	Aloe	0,052	0,105	1,9	27	96	3
<b>2</b>	<b>Baby1</b>	<b>0,064</b>	<b>0,101</b>	<b>1,1</b>	<b>30</b>	<b>95</b>	<b>5</b>
3	Baby2	0,032	0,101	0,3	31	96	4
4	Baby3	0,046	0,101	1,3	42	98	2
5	Bowling1	0,024	0,101	0,3	44	98	2
6	Bowling2	0,072	0,101	0,1	49	98	2
7	Cloth1	0,072	0,115	3,7	39	97	2
8	Cloth2	0,100	0,101	0,7	53	98	2
9	Cloth3	0,032	0,109	1,9	37	97	2
10	Cloth4	0,088	0,101	1,9	48	98	1
11	Flowerpots	0,042	0,101	0,5	48	98	2
12	Lampshade1	0,040	0,111	0,1	40	98	2
13	Lampshade2	0,025	0,105	0,2	42	98	2
14	Midd1	0,050	0,101	2,5	36	99	1
15	Midd2	0,027	0,123	2,5	33	99	1
16	Monopoly	0,078	0,129	0,9	28	97	2



17	Plastic	0,018	0,101	0,7	47	99	1
18	Rocks1	0,100	0,103	0,5	38	97	2
19	Rocks2	0,100	0,101	1,1	38	97	2
20	Wood1	0,074	0,101	0,5	42	98	1
21	Wood2	0,018	0,101	0,5	51	99	1

Analizując cechy dla minimalnych wartości  $\delta_D$  formułuje się następujące wnioski:

- dla obrazowań zawierających drobne elementy wartości  $p_{min}$  i  $\sigma$  zwiększyły swoje wartości w stosunku do stanu występującego dla minimalnych wartości  $\delta_r$ , co najmniej trzy krotnie,
- dla obrazowań, na których występują dominującą wielkością jednobarwne elementy wartości  $p_{min}$  i  $\sigma$  zmniejszyły swoje wartości w stosunku do stanu występującego dla minimalnych wartości  $\delta_r$ , co najmniej dwu krotnie.

Tab. 8.28 Minimalne wartości  $\delta_p$  dla operatora Canny'ego.

Lp	Nazwa	$p_{min}$	$p_{max}$	$\sigma$	$\delta_r$	$\delta_D$	$\Delta_r$
<b>1</b>	<b>Aloe</b>	<b>0,098</b>	<b>0,259</b>	<b>0,5</b>	<b>22</b>	<b>86</b>	13
2	Baby1	0,048	0,103	0,9	22	74	13
3	Baby2	0,016	0,121	0,1	24	77	48
4	Baby3	0,046	0,101	0,5	33	85	19
5	Bowling1	0,014	0,103	0,9	39	73	11
6	Bowling2	0,030	0,303	0,1	33	84	10
7	Cloth1	0,084	0,119	1,7	31	90	2
8	Cloth2	0,048	0,115	2,1	42	85	<b>49</b>
9	Cloth3	0,064	0,191	0,5	28	86	12
10	Cloth4	0,096	0,341	0,3	38	92	13
11	Flowerpots	0,020	0,101	0,1	40	85	20
12	Lampshade1	0,018	0,130	0,1	38	74	13
13	Lampshade2	0,016	0,121	0,7	35	76	21
14	Midd1	0,048	0,127	1,7	36	85	18
15	Midd2	0,040	0,101	0,9	36	87	29
16	Monopoly	0,010	0,101	0,1	28	80	19
17	Plastic	0,012	0,101	0,7	25	72	5
18	Rocks1	0,052	0,265	0,2	30	88	9
19	Rocks2	0,068	0,117	0,9	30	87	14

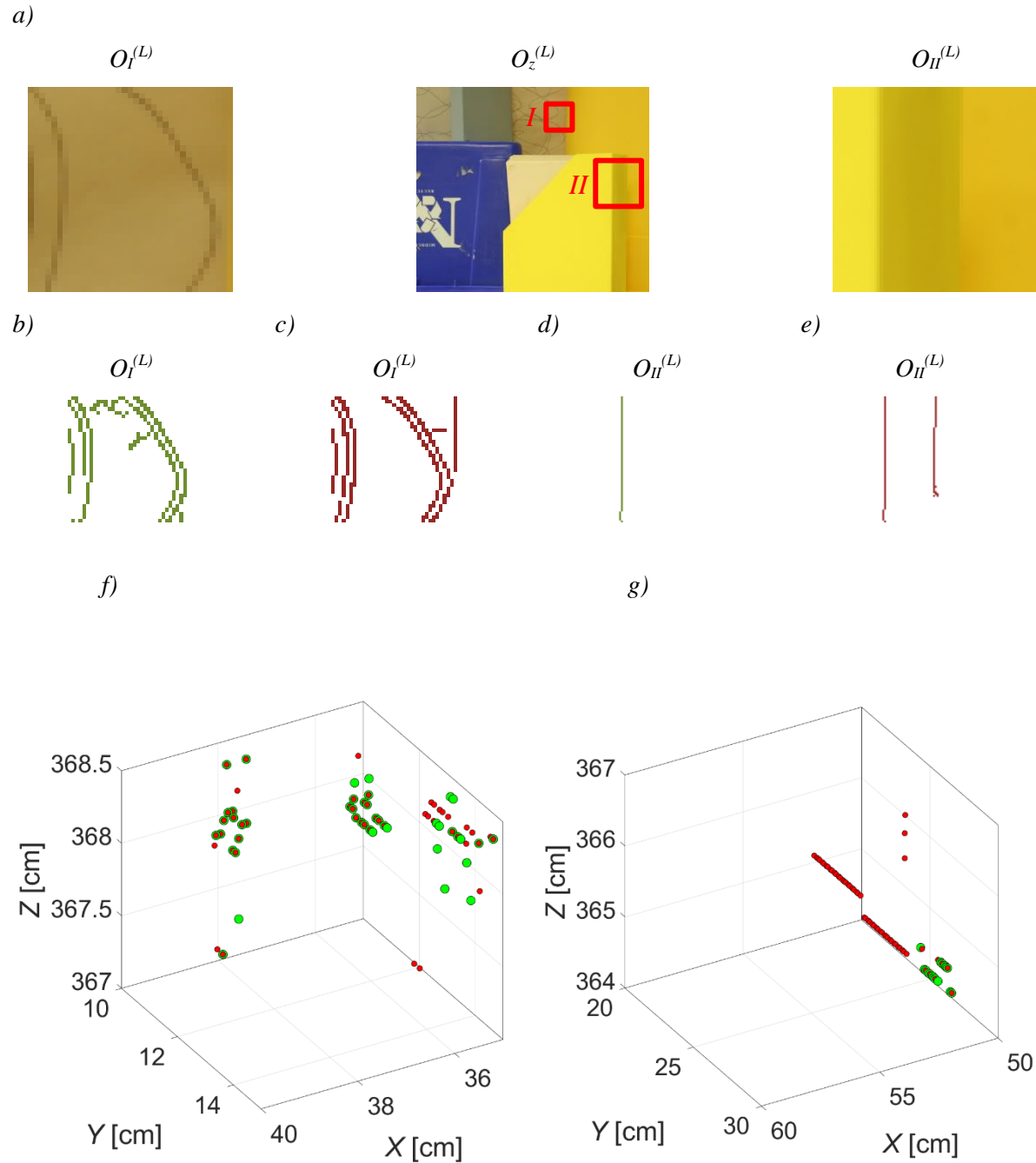
20	Wood1	0,024	0,311	0,3	32	88	33
21	Wood2	0,024	0,135	0,5	29	88	38

Na rys. 8.8 przedstawiono wynik rekonstrukcji przestrzennej dla cech pozwalających osiągnąć minimalne wartości błędów  $\delta_r$  i  $\delta_D$  dla przestrzeni Hough'a. Rekonstrukcję przeprowadzono dla dwóch fragmentów obrazu, charakterystycznych tym, iż obrazy binarne pierwszego fragmentu są do siebie zbliżone, zaś drugiego fragmentu można uznać za różne.

Tab. 8.29 Minimalne wartości  $\delta_{Dp}$  dla operatora Canny'ego.

Lp	Nazwa	$\rho_{min}$	$\rho_{max}$	$\sigma$	$\delta_r$	$\delta_D$	$\Delta_D$
1	Aloe	0,078	0,109	2,7	33	72	15
<b>2</b>	<b>Baby1</b>	<b>0,090</b>	<b>0,109</b>	<b>1,7</b>	<b>31</b>	<b>63</b>	15
3	Baby2	0,048	0,101	1,9	34	65	<b>19</b>
4	Baby3	0,100	0,103	1,5	44	76	11
5	Bowling1	0,068	0,101	0,5	48	69	14
6	Bowling2	0,100	0,121	0,7	37	77	10
7	Cloth1	0,046	0,167	2,7	33	83	11
8	Cloth2	0,098	0,101	1,3	44	84	12
9	Cloth3	0,086	0,105	2,5	34	71	18
10	Cloth4	0,094	0,101	2,1	43	85	10
11	Flowerpots	0,018	0,113	1,3	42	78	9
12	Lampshade1	0,046	0,159	0,5	44	69	8
13	Lampshade2	0,032	0,103	0,1	43	74	6
14	Midd1	0,052	0,101	2,1	38	82	9
15	Midd2	0,022	0,117	1,9	43	83	8
16	Monopoly	0,090	0,109	2,3	39	69	11
17	Plastic	0,018	0,101	0,7	25	71	2
18	Rocks1	0,024	0,111	2,5	33	76	15
19	Rocks2	0,100	0,101	1,9	34	74	17
20	Wood1	0,074	0,105	1,5	40	72	16
21	Wood2	0,054	0,109	0,2	32	87	4

Analiza wpływu cech na jakość odpowiedniości punktów dla  $\delta_{rp}$  i  $\delta_{Dp}$  pozwala wysunąć podobne wnioski co dla kryteriów  $\delta_r$  i  $\delta_D$ .



Rys. 8.8. Rekonstrukcja 3D metodą MED-NDD dwóch fragmentów obrazu Plastic oznaczonych jako obszar I i obszar II, wysegmentowanych operatorem Canny'ego przy zastosowaniu wartości cech minimalizujących błąd  $\delta_r$  (kolor zielony)(Tab. 8.30) i błąd  $\delta_D$  (kolor czerwony)(Tab. 8.31): a) obraz wejściowy z zaznaczonymi obszarami I i II; b) i d) wysegmentowane obszary dla cech minimalizujących  $\delta_r$ ; c) i e) wysegmentowane obszary dla cech minimalizujących  $\delta_D$ ; f) wynik rekonstrukcji 3D obszaru I jednocześnie dla cech minimalizujących  $\delta_r$  (kolor zielony) i  $\delta_D$  (kolor czerwony); g) wynik rekonstrukcji 3D obszaru II jednocześnie dla cech minimalizujących  $\delta_r$  (kolor zielony) i  $\delta_D$  (kolor czerwony).

W następnym podrozdziale podjęta zostanie próba całościowej analizy ilościowej, jak i jakościowej, otrzymanych wyników dla wszystkich pięciu przestrzeni. Zostaną również zwizualizowane i omówione najbardziej korzystne zestawy cech.

## 8.6. Podsumowanie analizy cech

Bazując na otrzymanych wynikach badań, zebrano minimalne wartości błędów skuteczności parowania i procentowej liczby błędnych sparowań oraz odpowiadające im cechy w tab. 8.30, tab. 8.31, tab. 8.32 i tab. 8.33.

Tab. 8.30 Cechy dla minimalnych wartości  $\delta_r$ .

przestrzeń	$\delta_r$	$\delta_D$	Optymalne cechy
<i>RGB</i>	25	95	$p_{RGB}=0,95, k_R=0,3, k_G=0,4, k_B=0,7$
<i>HSV</i>	26	100	$p_{HSV}=0,75, k_H=0,6, k_S=0,0, k_V=0,5$
<i>Lab</i>	26	100	$p_{Lab}=0,80, k_L=0,6, k_a=0,0, k_b=0,2$
Hough	9	39	$l_{max}=5-100, d_{prz}=5-100, d_{min}=1-63$
Canny	25	97	$p_{min}=0,036, p_{max}=0,101, \sigma=0,1$

Tab. 8.31 Cechy dla minimalnych wartości  $\delta_D$ .

przestrzeń	$\delta_D$	$\delta_r$	Optymalne cechy
<i>RGB</i>	95	26	$p_{RGB}=0,95, k_R=0,4, k_G=0,0, k_B=0,0$
<i>HSV</i>	95	28	$p_{HSV}=0,95, k_H=0,8, k_S=0,0, k_V=0,3$
<i>Lab</i>	95	26	$p_{Lab}=0,95, k_L=0,1-1,0, k_a=0,0, k_b=0,0$
Hough	39	9	$l_{max}=5-100, d_{prz}=5-100, d_{min}=1-63$
Canny	95	30	$p_{min}=0,064, p_{max}=0,101, \sigma=1,1$

Tab. 8.32 Cechy dla minimalnych wartości  $\delta_{rp}$ .

przestrzeń	$\delta_{rp}$	$\delta_{Dp}$	Optymalne cechy
<i>RGB</i>	28	70	$p_{RGB}=0,95, k_R=0,8, k_G=0,4, k_B=0,0$
<i>HSV</i>	24	78	$p_{HSV}=0,95, k_H=0,0, k_S=0,0, k_V=0,8$
<i>Lab</i>	25	79	$p_{Lab}=0,95, k_L=0,1-1,0, k_a=0,0, k_b=0,0$
Hough	12	55	$l_{max}=7, d_{prz}=20-30, d_{min}=1$
Canny	22	86	$p_{min}=0,098, p_{max}=0,259, \sigma=0,5$

Tab. 8.33 Cechy dla minimalnych wartości  $\delta_D$ .

przestrzeń	$\delta_D$	$\delta_r$	Optymalne cechy
<i>RGB</i>	58	36	$p_{RGB}=0,95, k_R=0,0, k_G=0,4, k_B=0,2$
<i>HSV</i>	59	38	$p_{HSV}=0,95, k_H=1,0, k_S=0,1, k_V=0,5$
<i>Lab</i>	56	38	$p_{Lab}=0,95, k_L=1,0, k_a=0,3, k_b=0,0$
Hough	52	13	$l_{max}=26-100, d_{prz}=5-12, d_{min}=35$
Canny	63	31	$p_{min}=0,090, p_{max}=0,109, \sigma=1,7$

Najwyższe wartości  $\Delta_r$  i  $\Delta_D$  dla poszczególnych kryteriów dokładności przedstawione w tab. 8.34, świadczą o istotnym wpływie cech na wyniki parowania.

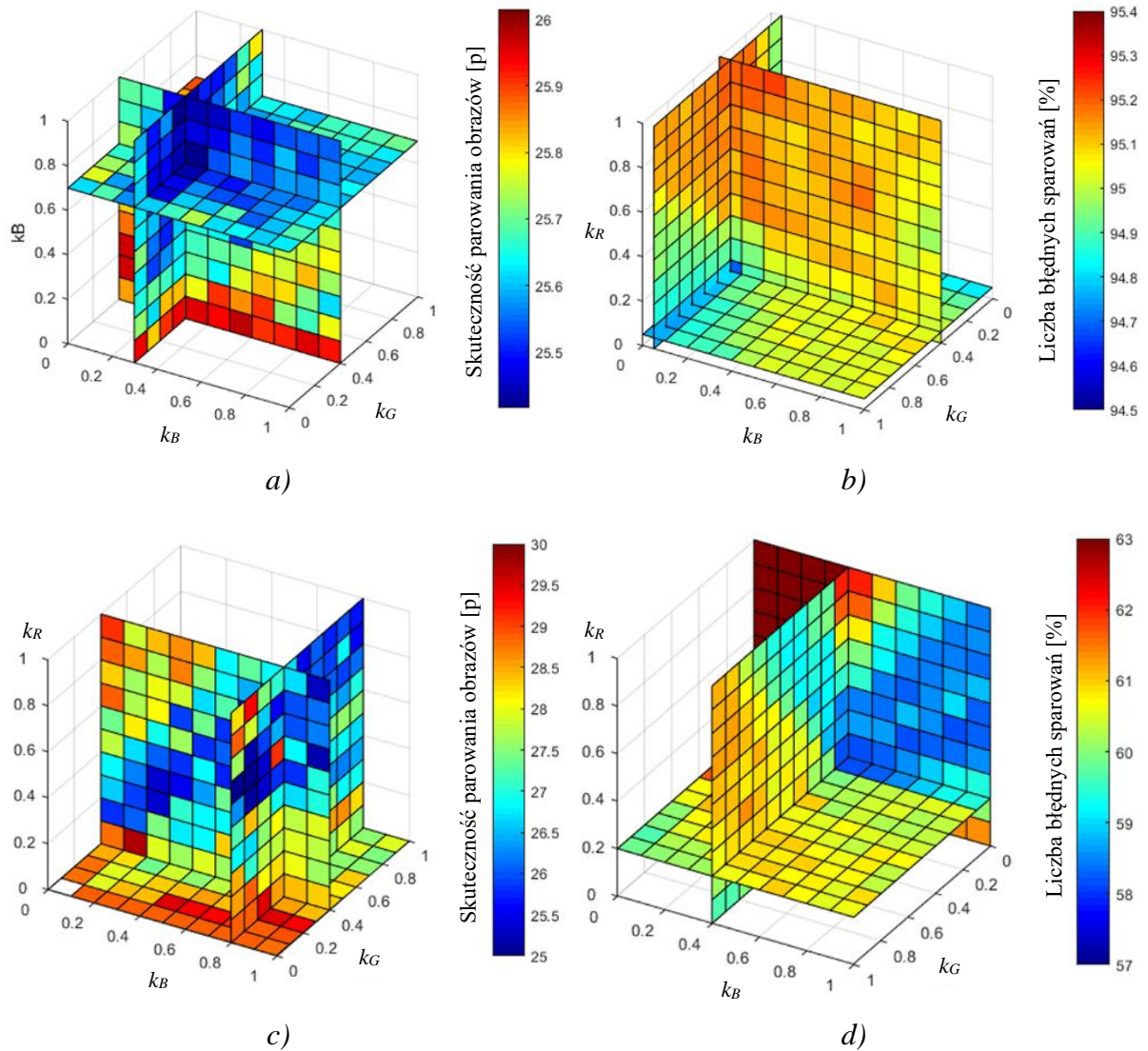
Tab. 8.34 Najwyższe i średnie wartości  $\Delta_r$  i  $\Delta_D$  dla poszczególnych przestrzeni.

przestrzeń	$\Delta_r \max$	$\bar{\Delta}_r$	$\Delta_D \max$	$\bar{\Delta}_D$
<i>RGB</i>	54	22	28	14
<i>HSV</i>	51	23	27	13
<i>Lab</i>	51	15	22	9
Hough	352	328	61	59
Canny	49	19	19	11

Zarówno dla  $\delta_r$ , jak i  $\delta_D$  maksymalny wpływ cech na wyniki parowania jest znaczący i w miarę jednolity, na poziomie  $\Delta_r \max = 50$  pikseli i  $\Delta_D \max = 20$  %. Jedynie dla przestrzeni Hough'a można zaobserwować znacznie większy skok różnicy o 352 piksele, co jest następstwem ustalenia zbyt szerokiego zakresu wartości cech w procedurze wykrywajProste( $l_{max}, d_{max}, d_{min}$ ), mimo, że zakresy te były wstępnie justowane.

Również średnie wartości  $\bar{\Delta}_r$  i  $\bar{\Delta}_D$  dla wszystkich przestrzeni potwierdzają znaczący wpływ zmian wartości cech na wyniki parowania.

Analiza trójwymiarowych wykresów najbardziej korzystnych zestawów cech dostarcza dalszych informacji o wpływie kontroli cech na wyniki poszukiwania odpowiedniości. Najbardziej korzystne zestawy cech, to cechy, dla których otrzymuje się minimalne wartości błędów. Podczas analizy wykresów przyjęto definicję płaszczyzn układu współrzędnych bazujących na koplanarnych osiach układu. Zdefiniowanie płaszczyzny jako  $k_R k_G$  oznacza, że płaszczyzna przechodzi przez osie  $k_R$  i  $k_G$ .



Rys. 8.9. Wizualizacja najkorzystniejszych cech dla przestrzeni  $RGB$  dla progu 0,95: a) dla błędów  $\delta_\delta$ ; b) dla błędów  $\delta_D$ ; c) dla błędów  $\delta_p$ ; d) dla błędów  $\delta_{Dp}$ .

Dla przestrzeni  $RGB$  na rys. 8.9 na wykresach a), b) i d) płaszczyzny cech przecinają się w punktach o najniższych wartościach błędów, co świadczy o poprawnym doborze cech. Dla wykresu a) najwyższe wartości błędu grupują się głównie na obrzeżach zbioru wyników w okolicach płaszczyzny  $k_G k_B$  dla wartości  $k_R = 0$ .

Wykres b) zdominowany jest przez wysokie wartości błędu. Najniższe wartości skupiają się na płaszczyźnie  $k_G k_B$  dla wartości  $k_R = 0$ ,  $k_G = 0$  i  $k_B = 0$ .

Na wykresie c) uwidacznia się płaszczyzna najmniejszego błędu w kolorze niebieskim przebiegająca od wartości  $k_R = 0.1$  pod kątem  $45^\circ$  do płaszczyzny  $k_R k_G$ , znajdująca się poza płaszczyznami optymalnych wartości cech wyznaczonych podczas badań. Przestrzeń wykresu, dla zapewnienia ciągłości i podobieństwa do pozostałych wykresów, odpowiada pełnemu

zbiorowi wyników dla progu 0,95. Ponieważ jest to wykres dla błędu  $\delta_p$ , gdzie jest on liczony dla  $D > 0$ , i wybierany z podzbioru wyników utworzonego przez macierze dysparycji o liczbie pikseli nie różniący się o więcej niż 30 % od  $l_c$ . Wykres ten potwierdza potrzebę zdefiniowania błędów  $\delta_p$  i  $\delta_{dp}$ . Dla niewielkiej liczby pikseli błędy mogą mieć zwodniczo niskie wartości, choć sama macierz dysparycji  $D$  jest ułomna z powodu niewielkiej liczby dysparycji, nie zapewniającej podobieństwa zrekonstruowanej sceny do oryginału. Najwyższe wartości błędu grupują się głównie na obrzeżach zbioru wyników.

Na wykresie *d*) widać wyraźny podział płaszczyzny  $k_R k_B$  w wartości  $k_G = 0$  przez płaszczyznę cechy  $k_B = 0,4$  na dwie części o diametralnie różnych wartościach błędu. Część płaszczyzny w stronę  $k_B = 1$  grupują się wartości najniższe błędu, podczas gdy w stronę  $k_B = 0$  wyraźnie grupują się wartości najwyższe błędu.

Znajomość wartości cech dla przestrzeni *RGB* pozwala zwiększyć dokładność parowania obrazów o maksymalnie 6 piksel dla  $\delta_r$  i 28 % dla  $\delta_D$ .

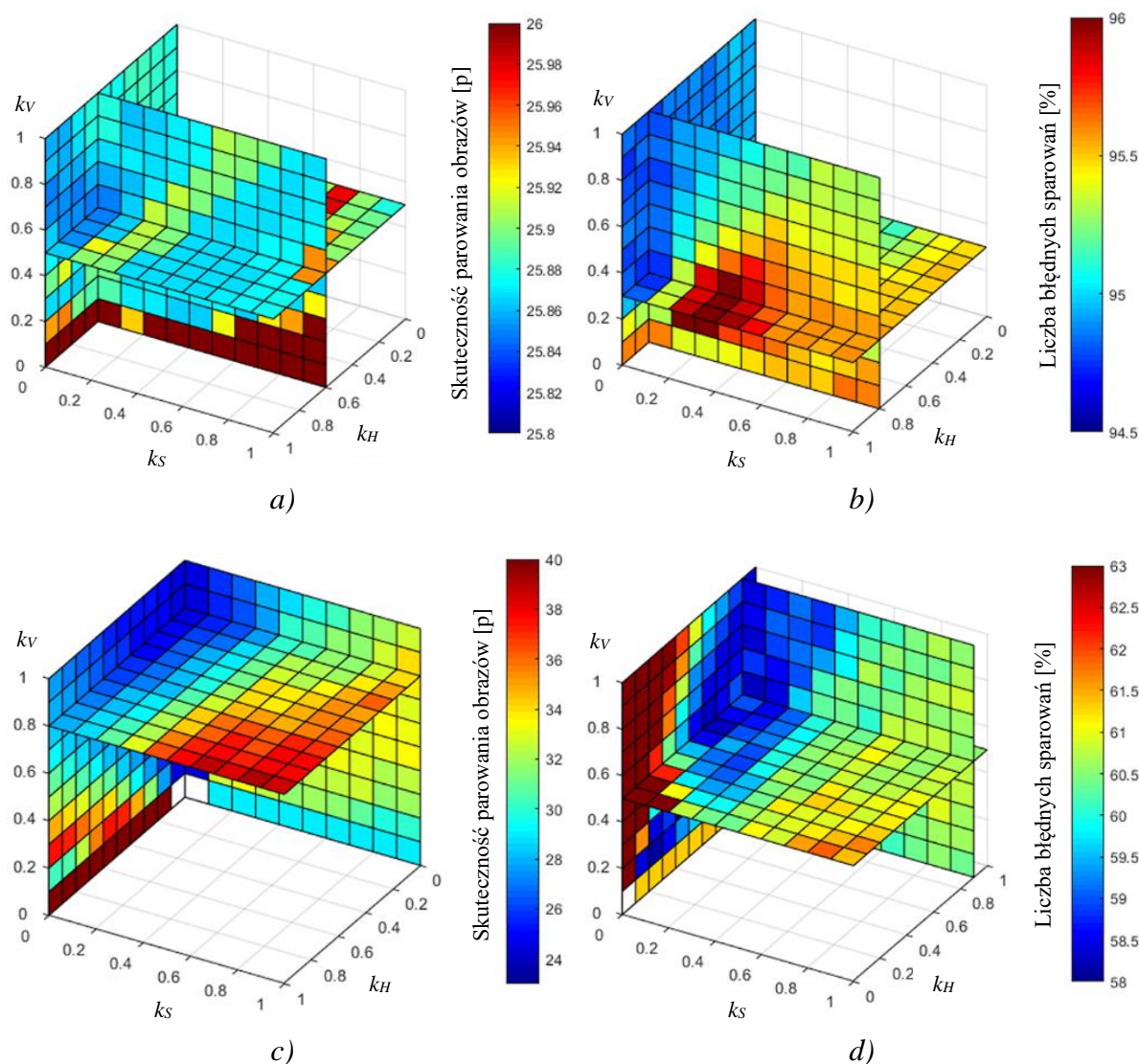
Dla wszystkich czterech wykresów na rys. 8.10 przedstawiających przestrzeń *HSV* płaszczyzny wyznaczane przez cechy przecinają się w punktach o najniższych wartościach błędów. Na wykresie *a*) najwyższe wartości błędów grupują się na płaszczyźnie  $k_H k_S$  dla wartości  $k_V = 0$ .

Dla wykresu *b*) najniższe wartości błędu grupują się powyżej ukośnej płaszczyzny położonej pod kątem  $45^\circ$  do płaszczyzny  $k_H k_V$  i powyżej prostej  $k_V = 0,2$ . Uwidacznia się kuliste ognisko najwyższych wartości błędów wokół punktu  $k_H = 0,8$ ,  $k_S = 0,3$ ,  $k_V = 0,3$ .

Dla wykresu *c*) najniższe wartości błędu grupują się wokół osi  $k_V$ , tworząc ostrosłup z wierzchołkiem w punkcie  $k_H=0, k_S=0, k_V=0$ . Najwyższe wartości błędu grupują się wokół osi  $k_H$  oraz wokół punktu  $k_H = 1$ ,  $k_S = 1$ ,  $k_V = 0,8$ .

Na wykresie *d*) można zaobserwować skupisko najniższych wartości błędów wokół punktu przecięcia się płaszczyzn wyznaczanych przez cechy z silnym trendem wznoszącym do wartości  $k_V = 1$ . Skupisko wartości najwyższych błędu występuje wokół punktu  $k_H = 0$ ,  $k_S = 0$ ,  $k_V = 1$  z silniejszym trendem opadającym ku płaszczyźnie  $k_H k_S$ .

Znajomość wartości cech dla przestrzeni *HSV* pozwala zwiększyć dokładność parowania obrazów o 20 piksel dla  $\delta_r$  i 27 % dla  $\delta_D$ .



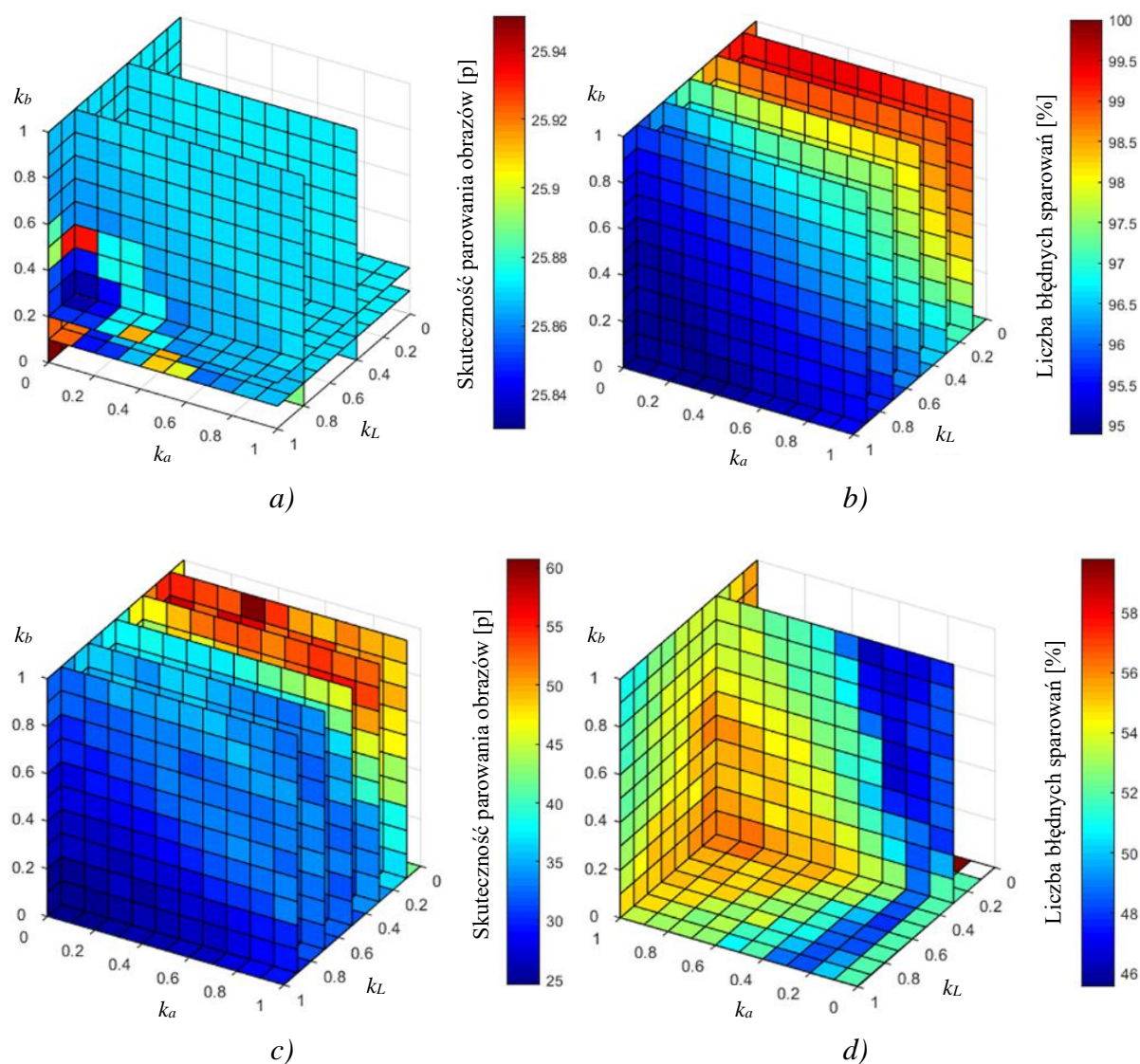
Rys. 8.10. Wizualizacja najkorzystniejszych cech dla przestrzeni HSV: a) dla błędów  $\delta_r$  dla progu 0,75; b) dla błędów  $\delta_D$  dla progu 0,95; c) dla błędów  $\delta_r$  dla progu 0,95; d) dla błędów  $\delta_{Dp}$  dla progu 0,95.

Na wykresie a) na rys. 8.11 przedstawiającym rozmieszczenie cech dla przestrzeni Lab najniższe wartości silnie skupiają się wokół punktu przecięcia płaszczyzn wyznaczonych przez cechy. Wyższe wartości błędów rozrzucone są punktowo bez wyraźnego trendu.

Na wykresie b) widoczne jest równomierne rozłożenie błędów. Minimalne wartości występują wokół punktu  $k_L = 1, k_a = 0, k_b = 0$  i rozciągają się wzdłuż osi  $k_L$ . Najwyższe wartości skupiają się wokół punktu  $k_L = 0,2, k_a = 1, k_b = 1$  i rozciągają się wzdłuż osi  $k_a$ .

Na wykresie c) najniższe wartości błędów skupiają się w punkcie  $k_L = 1, k_a = 0, k_b = 0$  i rozciągają się wzdłuż osi  $k_L$ , obejmują w całości dwie płaszczyzny  $k_L=0,9$  i  $k_L=0,7$ . Najwyższe wartości błędów grupują się wzdłuż osi  $k_a$  i dla wartości  $k_L = 0,3$  oraz  $k_b = 1$ .





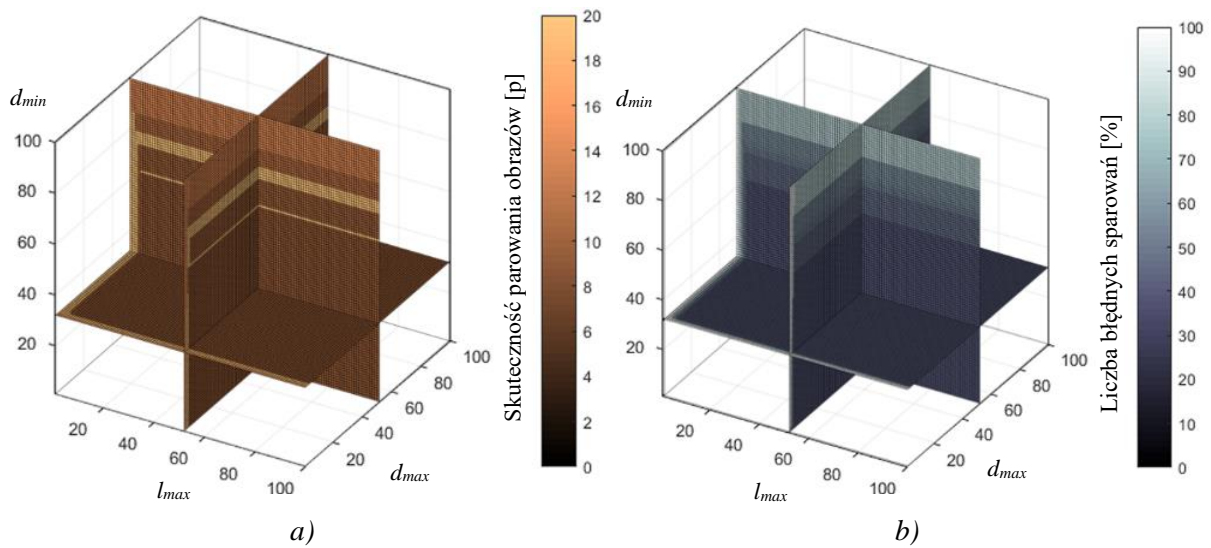
Rys. 8.11. Wizualizacja najkorzystniejszych cech dla przestrzeni *Lab*: a) dla błędów  $\delta_r$  dla progu 0,80; b) dla błędów  $\delta_D$  dla progu 0,95; c) dla błędów  $\delta_{rp}$  dla progu 0,95; d) dla błędów  $\delta_{Dp}$  dla progu 0,95.

Na wykresie d) silnie rysuje się niebieski obszar najmniejszego błędu zamknięty od dołu przez ukośną płaszczyznę położoną pod kątem  $20^\circ$  do płaszczyzny  $k_L k_b$ . Przyczyna wystąpienia tego obszaru poza obszarem wyznaczonym przez przecięcie się płaszczyzn cech jest dokładnie taki sam jak dla wykresu c) na rys. 8.9 dla przypadku przestrzeni *RGB*. Niewielka liczba pikseli na obrazach generuje niewielkie błędy, jednocześnie nie przystając do obrazu źródłowego. Dlatego wymagane jest przesianie wyników zgodnie z kryterium (7.3) uwzględniającym liczbę Canny'ego  $l_c$ .

Znajomość wartości cech dla przestrzeni *Lab* pozwala zwiększyć dokładność parowania obrazów o 9 pikseli dla  $\delta_r$  i 5 % dla  $\delta_D$ .

Dla przestrzeni Hough'a wartości błędów przedstawione na rys. 8.12, grupują się w wyraźnie od siebie oddzielone, warstwowo ułożone obszary. Na wykresie *a)* przewaga najniższych wartości błędów grupuje się w dolnej warstwie. Powyżej niej na przemian układają się warstwy wyższych i niższych wartości błędów. Ten sam scenariusz obserwujemy na wykresie *b)*.

Znajomość wartości cech dla przestrzeni Hough'a pozwala zwiększyć dokładność parowania obrazów o 304 piksele dla  $\delta_r$  i 61 % dla  $\delta_D$



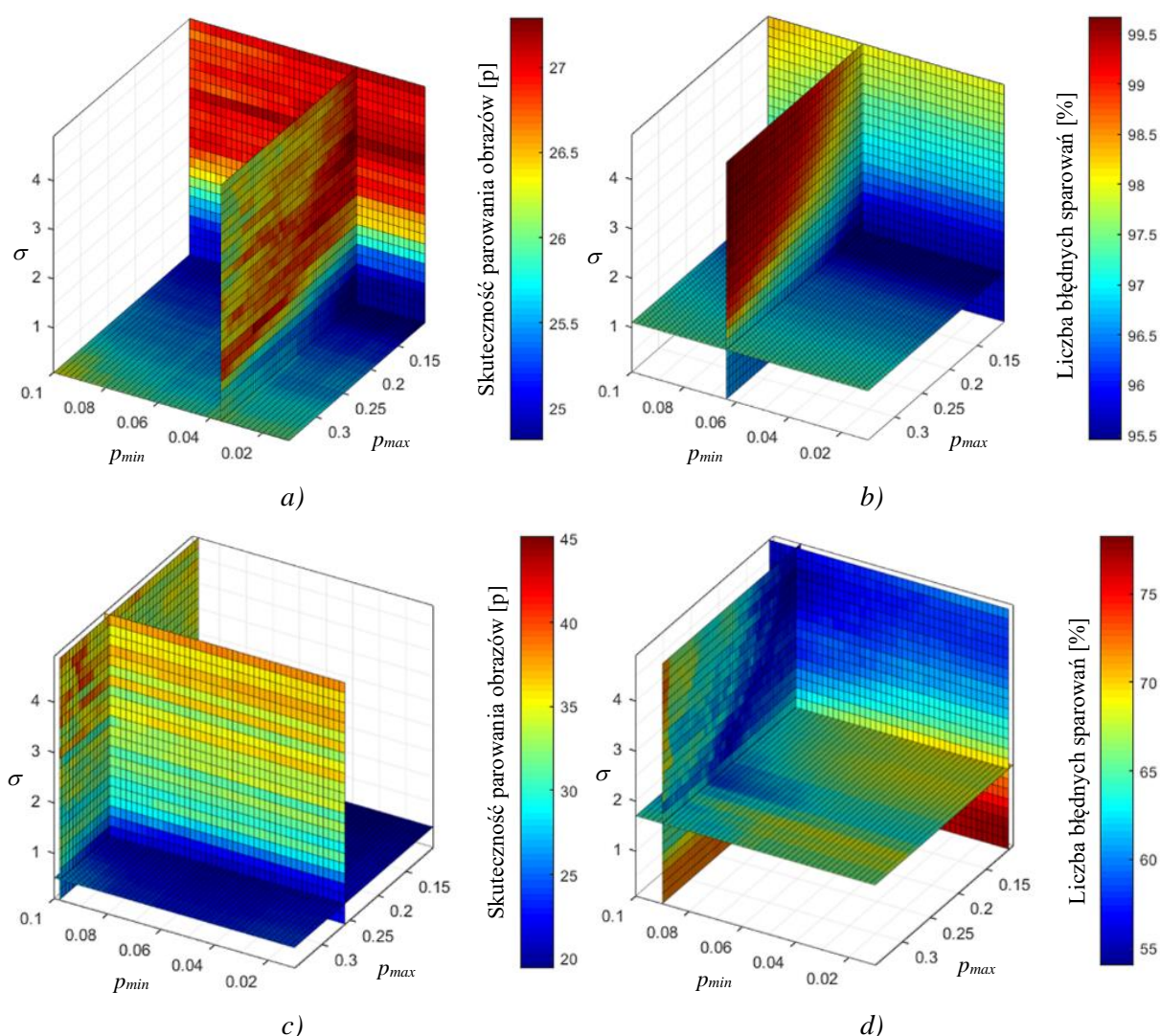
Rys. 8.12. Wizualizacja najkorzystniejszych cech dla przestrzeni Hough'a: *a)* dla błędów  $\delta_r$ ; *b)* dla błędów  $\delta_D$ .

Na wykresie *a)* przedstawionym na rys. 8.13 wartości najniższych błędów grupują się wzdłuż osi  $p_{min}$ . Wartości najwyższe błędów grupują się w górnej części płaszczyzny  $p_{min}\sigma$ , wchodząc klinem w zbiór rozwiązań.

Na wykresie *b)* najniższe wartości błędów ponownie grupują się wzdłuż osi  $p_{min}$ . Najwyższe wartości błędów grupują się po przeciwnej stronie zbioru rozwiązań, wzdłuż osi  $p_{min}$  dla  $p_{max} = 0,35$  i  $\sigma = 5$ .

Na wykresie *c)* najniższe wartości grupują się na płaszczyźnie  $p_{min}p_{max}$  poniżej  $\sigma = 0,6$ . Wartości błędów równomiernie rosną w górę wykresu, osiągają odosobnione wartości maksymalne.

Na wykresie *d)* po przesianiu zbioru wyników w poszukiwaniu wartości  $l_p$  zbliżonych do  $l_c$ , płaszczyzny wyznaczone przez wartości cech przecinają się poza zgrupowaniem wartości minimalnych błędów, które grupują się wzdłuż osi  $p_{min}$  dla  $\sigma = 5$ . Wartości maksymalne błędów grupują się zaraz poniżej zgrupowania wartości najniższych wzdłuż osi  $p_{min}$ .



Rys. 8.13. Wizualizacja najkorzystniejszych cech dla przestrzeni operatora Canny'ego: a) dla błędów  $\delta_r$ ; b) dla błędów  $\delta_D$ ; c) dla błędów  $\delta_p$ ; d) dla błędów  $\delta_{Dp}$ .

Znajomość wartości cech dla przestrzeni operatora Canny'ego pozwala zwiększyć dokładność parowania obrazów o 13 pikseli dla  $\delta_r$  i 15 % dla  $\delta_D$ . Dla większości wykresów obserwujemy stopniowy spadek wartości  $\delta_r$  i  $\delta_D$ , aż do punktu skrzyżowania się wszystkich trzech płaszczyzn, gdzie dominuje najciemniejszy kolor, symbolizujący najniższe wartości. Wyjątkiem jest przedstawienie przestrzeni Hough'a, gdzie zmiany zachodzą zakresami, co dobrze ilustrują wykresy na rys. 8.12.

Podsumowując badania można wysunąć następujące wnioski:

- zmieniając wartość cech znaczących wpływa się w znaczny sposób na dokładność wyników parowania obrazów, co wynika z tab. 8.34 i co świadczy o właściwym wyborze cech znaczących,

- błąd skuteczności parowania obrazów  $\delta_r$  można poprawić o maksymalnie 54 i średnio o 20 piksele, a w przypadku przestrzeni Hough'a nawet 352 piksele,
- błąd procentowej liczby błędnych sparowań  $\delta_D$  można poprawić o maksymalnie 28 % i średnio 12 %, a w przypadku przestrzeni Hough'a nawet o 61 %,
- analiza cech w przestrzeni *RGB*, *HSV* i *Lab* wykazuje wyróżnianie znaczącymi wartościami cech kanałów o wysokiej kontrastowości,
- różnorodność zobrazowań pod względem radiometrycznym, jak i zawartości pod względem wielkości obiektów tworzących scenę, ich rozmieszczenia na obrazie i wzajemnego układu, uniemożliwia wysunięcie dalszych wniosków, co do preferowanych wartości cech.

By wysunąć dalej idące wnioski należałoby przeprowadzić ponowne badania na specjalnie do tego celu sprofilowanych obrazach. Opierając się na dotychczas pozyskanych wynikach, można wysunąć wniosek, iż należałoby przygotować zobrazowania wyraźnie zróżnicowane pod względem radiometrycznym, zawierające kilka grup z dominującą daną kolorystyką oraz znacznymi różnicami w oświetleniu sceny. Następnie należałoby przygotować kilka grup obrazów wypełnionych małymi, średnimi i dużymi obiektami, o gładkiej i zróżnicowanej fakturze oraz wyraźnym dającym się sklasyfikować rozmieszczeniu, zarówno na obrazie, jak i względem siebie. Przeprowadzenie analiz na tak przygotowanym materiale badawczym mogłoby pozwolić na wysunięcie wniosków dotyczących uniwersalnych wartości cech zapewniających maksymalną efektywność.

## 9. DYSKUSJA

Poniżej przedstawione **znane metody** poszukiwania odpowiedniości dla punktów na obrazach stereowizyjnych stanowią bazę porównawczą dla metod opracowanych w niniejszej pracy.

- W pracy [34] zaproponowano wariant metody opartej o znormalizowaną korelację krzyżową (ang. *normalised cross correlation* - NCC). w [34] przedstawiono oryginalne dane z publikacji [34].

Tab. 9.35 Porównanie skuteczności działania zaproponowanego przez Lewis'a algorytmu Fast NCC z istniejącym Direct NCC.

rozmiar okna	długość sekwencji	Direct NCC	Fast NCC
168×86	896 klatek	900 min.	102 min.
115×200, 150×150	490 klatek	858 min.	57 min.

Pomiary przeprowadzono na komputerze klasy PC z procesorem R4000 taktowanym zegarem 100 MHz.

- W pracy [37] przedstawiono metodę opartą o rachunek wariacyjny (VAR). w [37] przedstawiono oryginalne wyniki porównania metody VAR do innych metod pod względem liczby błędnych sparowań.

Tab. 9.36 Porównanie skuteczności działania algorytmu VAR do innych metod na przykładzie procentowej liczby błędnych sparowań  $\delta_D$  dla progu  $p_D = 1$  piksel. Kolejność metod zgodna z oryginalną publikacją [37].

Algorytm	Tsukuba	Venus	Teddy	Cones
DoubleBP [81]	1.24	0.35	4.13	7.02
CoopRegion [82]	1.00	0.18	3.41	5.95
VAR	3.13	1.07	2.30	3.43
MultiResGC [83]	1.05	0.46	7.13	8.80
RealtimeBP [84]	3.04	1.53	8.27	9.11
RealtimeGPU [85]	3.27	1.90	8.65	10.80
Infection [86]	7.81	3.66	18.3	16.60



Pomiar został przeprowadzony w ujednoliconym systemie do badania sprawności algorytmów poszukiwania odpowiedniości udostępnionym na stronie internetowej Middlebury College.

- W pracy [45] zaproponowano metodę stochastyczną opartą o funkcję kosztu bazującą na informacji wzajemnej (SGBM), którą niosą ze sobą obrazy. Zaczerpnięte z publikacji oryginalne wyniki porównania liczby błędnych sparowań z innymi metodami przedstawiono w tab. 9.37.

Tab. 9.37 Porównanie skuteczności działania algorytmu SGBM z innymi metodami na przykładzie procentowej liczby błędnych sparowań  $\delta_D$  dla progu  $p_D = 0.5$  piksela. Kolejność metod zgodna z oryginalną publikacją [45].

Algorytm	Tsukuba	Venus	Teddy	Cones
C-SGBM	13.9	3.30	9.82	5.37
SGBM	13.4	4.55	11.00	4.93
AdaptingBP [87]	19.1	4.84	12.80	7.02
Segm+visib [88]	12.7	10.40	11.00	8.12
DoubleBP	18.7	7.85	14.30	11.9
GenModel [89]	7.89	4.59	14.80	10.2

Pomiar został, podobnie jak dla poprzedniej metody, przeprowadzony w ujednoliconym systemie do badania sprawności algorytmów poszukiwania odpowiedniości udostępnionym na stronie internetowej Middlebury College.

Przedstawione w Rozdziale 6 metody parowania obrazów binarnych, tj. MED, MED-NDD i MED-RGB, **opracowane przez autora** niniejszej rozprawy, porównano z przedstawionymi wyżej znanymi i najczęściej stosowanymi metodami parowania obrazów stereowizyjnych, tj. NCC, VAR i SGBM, uznawanymi za najbardziej skuteczne dla szerokiego spektrum zastosowań [45, 58, 90, 91].

Wykorzystywane wcześniej obrazy, tj. obrazy brył stereometrycznych, naczyń podskórnych oraz obrazów z biblioteki Middlebury poddano segmentacji metodą BGOM<sup>2</sup>,

<sup>2</sup> Parowanie obrazów Middlebury przez metody NCC, VAR i SGBM zostało wykonane w przestrzeni *RGB* z pominięciem segmentacji.

a następnie procesowi poszukiwania odpowiedniości istniejącymi metodami parowania obrazów NCC, VAR i SGBM oraz nowo zaproponowanymi metodami MED, MED-NDD i MED-RGB.

Porównanie wyników zrealizowano na bazie kryteriów przedstawionych w podrozdziale 7.1, tj.:

- czasu działania metod  $t$ ,
- skuteczności parowania  $\delta_r$ ,
- procentowej liczby błędnych sparowań  $\delta_D$ .

Metody VAR i SGBM zawdzięczają swoją szybkość zoptymalizowanemu kodowi, co nieco zaburza porównanie wyników badań. Zdecydowano się jednak na zamieszczenie porównań z udziałem tych algorytmów z powodu ich znaczenia i powszechności stosowania. Dla NCC zastosowano maskę 3 pikselową.

W następnym podrozdziale przedstawione zostanie porównanie czasów działania  $t$  istniejących i nowych algorytmów znajdowania odpowiedniości.

## 9.1. Czas parowania $t$ dla różnych obrazów

Porównanie czasu parowania  $t$  wykonano dla obrazów własnych, tj. brył stereometrycznych i naczyń podskórnych, oraz obrazów Middlebury i wyniki zamieszczono w tab. 9.38, tab. 9.39, tab. 9.40 oraz zwizualizowano na rys. 9.14, rys. 9.15. Obliczenia zrealizowano w środowisku Matlab na komputerze klasy PC z procesorem Intel Core i5 3,1 GHz i 8 GB pamięci RAM. Mierzony był jedynie czas działania algorytmu parowania obrazów z pominięciem wszelkich innych działań, jak wczytywanie obrazów, czy ich segmentacja. w tabelach wyłuszczone wartości minimalne.

Tab. 9.38 Przedstawienie czasów parowania  $t$  dla obrazów własnych.

	Pir	Cube	Ves1	Ves2	Ves3	Ves4	Ves5	Ves6	Ves7	Ves8	Ves9	Ves10	średnia
<b>MED</b>	<b>0,02</b>	<b>0,03</b>	<b>0,01</b>	<b>0,01</b>	<b>0,01</b>	<b>0,01</b>	<b>0,01</b>	<b>0,01</b>	<b>0,01</b>	<b>0,01</b>	<b>0,01</b>	<b>0,01</b>	<b>0,01</b>
<b>MED-NDD</b>	0,03	0,04	0,02	0,02	<b>0,01</b>	0,02	<b>0,01</b>	<b>0,01</b>	0,06	0,02	0,02	0,02	0,02
<b>MED-RGB</b>	0,03	0,07	0,03	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,03
NCC	27,7	26,88	80,14	78,22	76,81	76,94	83,23	77,45	77,31	78,00	114,01	118,19	76,24
VAR	0,55	0,55	1,33	1,30	1,27	1,27	1,31	1,28	1,29	1,29	1,78	1,83	1,25
SGBM	0,02	0,02	0,04	0,04	0,04	0,04	0,04	0,04	0,04	0,04	0,05	0,05	0,04

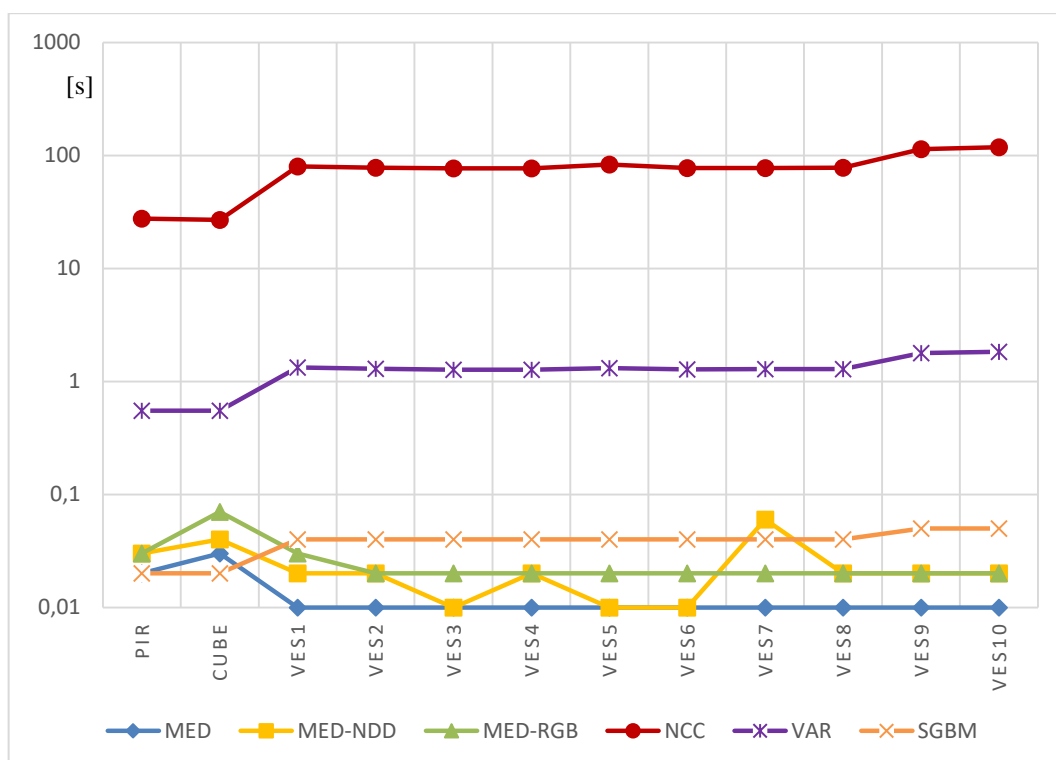
Tab. 9.39 Przedstawienie czasów parowania  $t$  dla obrazów Middlebury - część 1.

	Aloe	Baby1	Baby2	Baby3	Bowling1	Bowling2	Cloth1	Cloth2	Cloth3	Cloth4	Flowerpots	średnia
<b>MED</b>	<b>0,04</b>	<b>0,03</b>	<b>0,03</b>	<b>0,05</b>	<b>0,03</b>	<b>0,02</b>	<b>0,02</b>	<b>0,03</b>	<b>0,03</b>	<b>0,03</b>	<b>0,02</b>	<b>0,03</b>
<b>MED-NDD</b>	0,17	0,93	0,06	0,06	0,06	0,07	0,12	0,05	0,10	0,08	0,06	0,16
<b>MED-RGB</b>	0,51	2,16	0,19	0,15	0,13	0,14	0,47	0,09	0,18	0,16	0,10	0,39
NCC	167,81	156,34	158,62	168,05	158,95	176,54	158,69	165,89	157,69	165,79	167,06	163,77
VAR	3,70	4,67	3,66	3,91	3,63	4,54	3,85	3,93	3,67	3,96	4,08	3,96
SGBM	0,07	0,07	0,07	0,07	0,06	0,07	0,07	0,06	0,06	0,06	0,06	0,07

Tab. 9.40 Przedstawienie czasów parowania  $t$  dla obrazów Middlebury - część 2.

	Lampshade1	Lampshade2	Mid1	Mid2	Monopoly	Plastic	Rocks1	Rocks2	Wood1	Wood2	średnia
<b>MED</b>	<b>0,02</b>	<b>0,02</b>	<b>0,02</b>	<b>0,02</b>	<b>0,02</b>	<b>0,02</b>	<b>0,02</b>	<b>0,02</b>	<b>0,02</b>	<b>0,02</b>	<b>0,02</b>
<b>MED-NDD</b>	0,03	0,03	0,05	0,07	0,04	0,03	0,08	0,08	0,06	0,09	0,06
<b>MED-RGB</b>	0,04	0,04	0,12	0,17	0,10	0,04	0,16	0,18	0,10	0,20	0,12
NCC	166,24	164,97	179,84	175,93	171,09	160,95	161,66	162,81	175,73	166,66	168,59
VAR	4,00	3,97	4,35	4,02	4,03	3,78	3,83	3,82	4,17	3,91	3,99
SGBM	0,07	0,07	0,07	0,07	0,06	0,06	0,06	0,06	0,07	0,06	0,07

Pod względem czasu działania najlepiej prezentuje się proponowana metoda MED, osiągając czasy rzędu 10 ms dla zobrażeń naczyń i 20 ms dla modeli stereometrycznych. Nieznacznie dłuższe czasy działania osiągają algorytmy MED-NDD i MED-RGB.

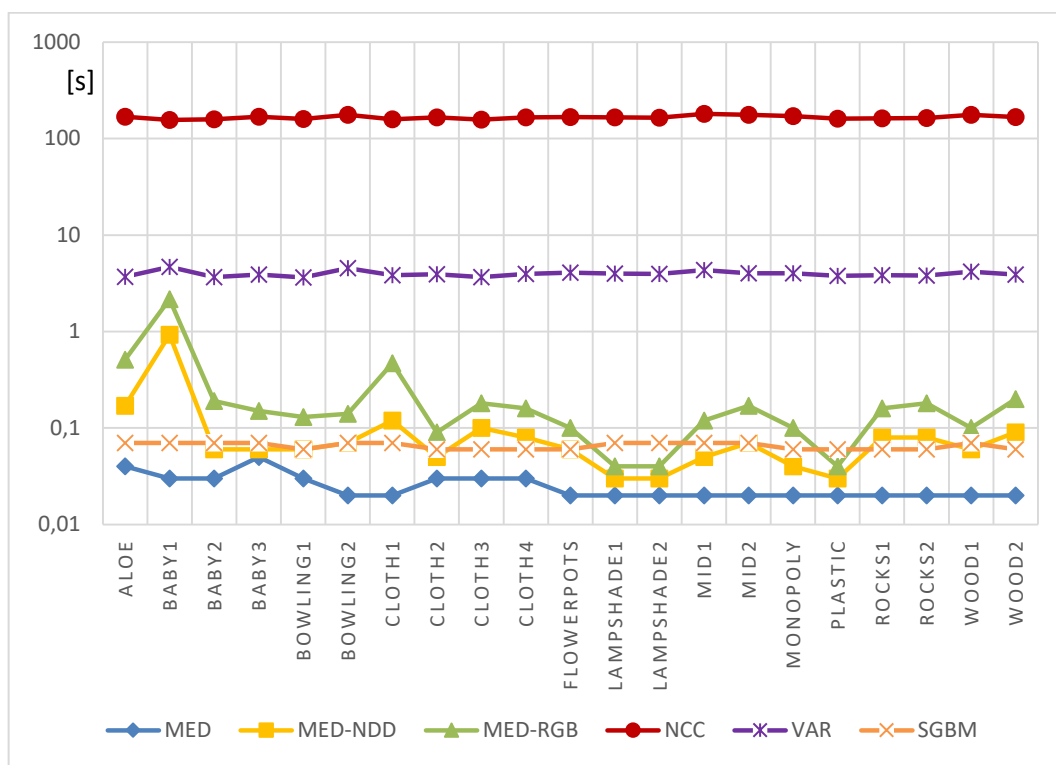


Rys. 9.14. Czas parowania  $t$  [s] dla obrazów własnych. Czas na osi rzędnych przedstawiono w skali logarytmicznej.

Algorytm MED parował obraz Ves10 prawie 12 tys. razy szybciej niż NCC, 183 razy szybciej niż VAR i 5 razy szybciej niż SGBM. Również z analizy wartości średnich w tabelach



wynika znaczna przewaga algorytmów proponowanych nad istniejącymi, powszechnie stosowanymi NCC, VAR i SGBM.



Rys. 9.15. Czas parowania  $t$  [s] dla obrazów Middlebury. Czas na osi rzędnych przedstawiono w skali logarytmicznej.

W następnym podrozdziale przedstawione zostanie porównanie skuteczności parowania  $\delta_r$  istniejących i nowych algorytmów znajdowania odpowiedniości.

## 9.2. Skuteczność parowania

Porównanie skuteczności parowania  $\delta_r$  dla wszystkich metod zamieszczono w tab. 9.41, tab. 9.42, tab. 9.43 oraz zwizualizowano na rys. 9.16, rys. 9.17. Skuteczność parowania  $\delta_r$  została przedstawiona w rozdziale 7.1. w tabelach wytłuszczono wartości minimalne.

Tab. 9.41 Przedstawienie skuteczności parowania  $\delta_r$  dla obrazów własnych.

	Pir	Cube	Ves1	Ves2	Ves3	Ves4	Ves5	Ves6	Ves7	Ves8	Ves9	Ves10	średnia
<b>MED</b>	13	8	3	<b>0</b>	3	2	<b>1</b>	8	2	4	<b>4</b>	<b>1</b>	4
<b>MED-NDD</b>	<b>6</b>	<b>5</b>	3	<b>0</b>	<b>2</b>	<b>1</b>	<b>1</b>	8	<b>1</b>	<b>3</b>	<b>4</b>	<b>1</b>	<b>3</b>
<b>MED-RGB</b>	9	6	<b>2</b>	<b>0</b>	3	2	<b>1</b>	<b>7</b>	<b>1</b>	4	<b>4</b>	<b>1</b>	<b>3</b>
NCC	29	29	20	30	30	30	20	30	30	30	20	19	26
VAR	30	30	20	30	30	30	20	30	30	30	20	20	27
SGBM	29	29	20	30	30	30	20	30	30	30	20	20	27

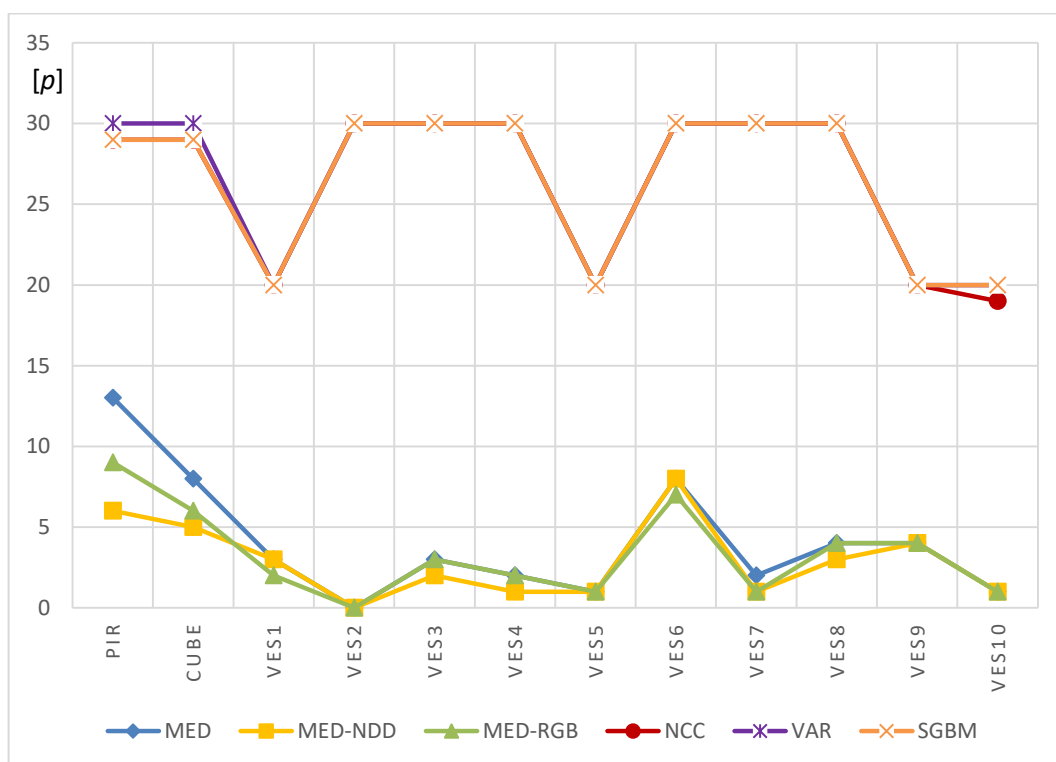
Tab. 9.42 Przedstawienie skuteczności parowania  $\delta_r$  dla obrazów Middlebury – część 1.

	Aloe	Baby1	Baby2	Baby3	Bowling1	Bowling2	Cloth1	Cloth2	Cloth3	Cloth4	Flowerpots	średnia
<b>MED</b>	18	15	10	21	21	26	21	28	26	21	27	21
<b>MED-NDD</b>	17	13	9	21	20	26	18	20	23	22	29	20
<b>MED-RGB</b>	16	16	11	21	23	24	19	21	22	21	30	20
NCC	14	14	14	16	24	22	6	22	10	16	22	16
VAR	<b>2</b>	<b>2</b>	<b>4</b>	<b>5</b>	12	13	<b>3</b>	<b>6</b>	<b>3</b>	<b>6</b>	<b>4</b>	<b>5</b>
SGBM	6	3	<b>4</b>	10	<b>8</b>	<b>10</b>	4	17	<b>3</b>	<b>6</b>	14	8

Tab. 9.43 Przedstawienie skuteczności parowania  $\delta_r$  dla obrazów Middlebury – część 2.

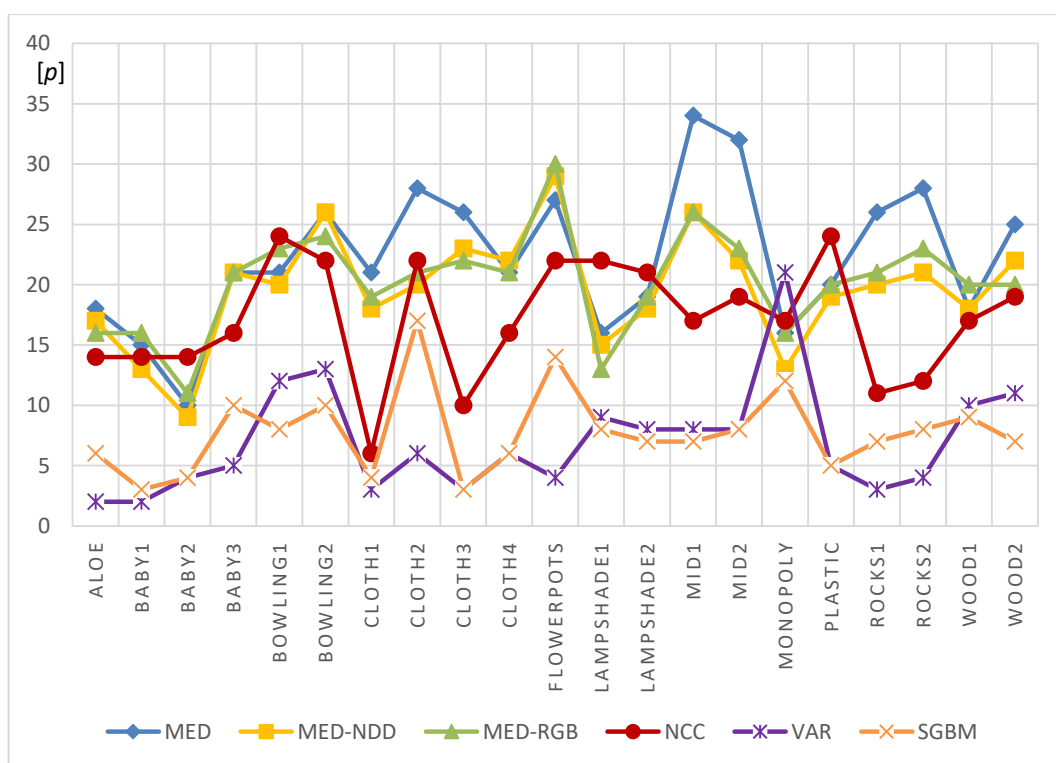
	Lampshade1	Lampshade2	Mid1	Mid2	Monopoly	Plastic	Rocks1	Rocks2	Wood1	Wood2	średnia
<b>MED</b>	16	19	34	32	16	20	26	28	18	25	23
<b>MED-NDD</b>	15	18	26	22	13	19	20	21	18	22	19
<b>MED-RGB</b>	13	19	26	23	16	20	21	23	20	20	20
NCC	22	21	17	19	17	24	11	12	17	19	18
VAR	9	8	8	<b>8</b>	21	<b>5</b>	<b>3</b>	<b>4</b>	10	11	9
SGBM	<b>8</b>	<b>7</b>	<b>7</b>	<b>8</b>	<b>12</b>	<b>5</b>	7	8	<b>9</b>	<b>7</b>	<b>8</b>

Dla obrazów naczyń skuteczność parowania  $\delta_r$  osiąga minimalną wartość bliską 0 dla obrazu Ves2 dla wszystkich trzech nowo proponowanych algorytmów. Dla reszty zobrazowań naczyń Ves1 – Ves10 wartości  $\delta_r$  nie przekraczają 8 pikseli, przy średniej wartości  $\delta_r$  dla algorytmu MED-NDD wynoszącej 2 piksele oraz 3 piksele dla MED i MED-RGB. Skuteczność parowania  $\delta_r$  dla istniejących metod dobija do wartości określonych przez maksymalne dysparycje, które w zależności od obrazów wynoszą 20 lub 30 pikseli.

Rys. 9.16. Skuteczność parowania  $\delta_r$  [p] dla obrazów własnych.

Dla przypadku modeli brył skuteczność parowania  $\delta_r$  nowych metod kształtuje się również znacznie lepiej od istniejących algorytmów NCC, VAR i SGBM. Najmniejszą wartość 5 pikseli osiąga metoda MED-NDD dla obrazu Cube, a największą wartość 30 pikseli osiąga algorytm VAR dla obu brył. Wartości  $\delta_r$  osiągane przez algorytmy NCC, VAR, SGBM dla brył stereometrycznych osiągają wartości zbliżone do wartości dysparycji dla tych obrazów.

Skuteczność parowania  $\delta_r$  dla obrazów Middlebury osiąga wartość najmniejszą 2 pikseli dla obrazów Aloe i Baby1 dla algorytmu VAR. Również w ogólnym ujęciu ten algorytm osiąga najlepsze wyniki w parowaniu obrazów Middlebury ze średnią wartością  $\delta_r$  wynoszącą 7 pikseli. Również algorytm SGBM wykazuje wysoką efektywność dla obrazów Middlebury ze średnią wartością  $\delta_r$  wynoszącą 8 pikseli.



Rys. 9.17. Skuteczność parowania  $\delta_r$  [p] dla obrazów Middlebury.

Wartość największą 34 pikseli osiąga metoda MED dla obrazu Mid1. Należy jednak zaznaczyć, iż nie jest to wartość zbliżona do największej wartości paralaksy dla tego obrazu, która wynosi 207 pikseli.

W następnym podrozdziale przedstawione zostanie porównanie liczby błędnych sparowań  $\delta_D$  dla istniejących i nowych algorytmów znajdowania odpowiedniości.

### 9.3. Liczba błędnych sparowań

Porównanie procentowej liczby błędnych sparowań zamieszczono w tab. 9.44, tab. 9.45, tab. 9.46 i zwizualizowano na rys. 9.18 i rys. 9.19. Definicja liczby błędnych sparowań  $\delta_D$  została przedstawiona w rozdziale 7.1. w tabelach wytłuszczono wartości minimalne.

Tab. 9.44 Przedstawienie procentowej liczby błędnych sparowań  $\delta_D$  dla obrazów własnych.

	Pir	Cube	Ves1	Ves2	Ves3	Ves4	Ves5	Ves6	Ves7	Ves8	Ves9	Ves10	średnia
<b>MED</b>	18	24	8	<b>0</b>	8	11	<b>2</b>	17	7	9	12	1	10
<b>MED-NDD</b>	<b>5</b>	16	7	<b>0</b>	<b>5</b>	<b>5</b>	<b>2</b>	17	<b>5</b>	7	<b>9</b>	<b>0</b>	<b>7</b>
<b>MED-RGB</b>	10	<b>14</b>	<b>4</b>	<b>0</b>	8	11	<b>2</b>	<b>14</b>	6	<b>6</b>	11	<b>0</b>	<b>7</b>
NCC	98	97	99	100	100	100	100	100	100	100	99	100	99
VAR	100	100	100	100	100	100	100	100	100	100	100	100	100
SGBM	98	97	99	100	100	99	99	100	100	99	99	100	99

Tab. 9.45 Przedstawienie procentowej liczby błędnych sparowań  $\delta_D$  dla obrazów Middlebury – część 1.

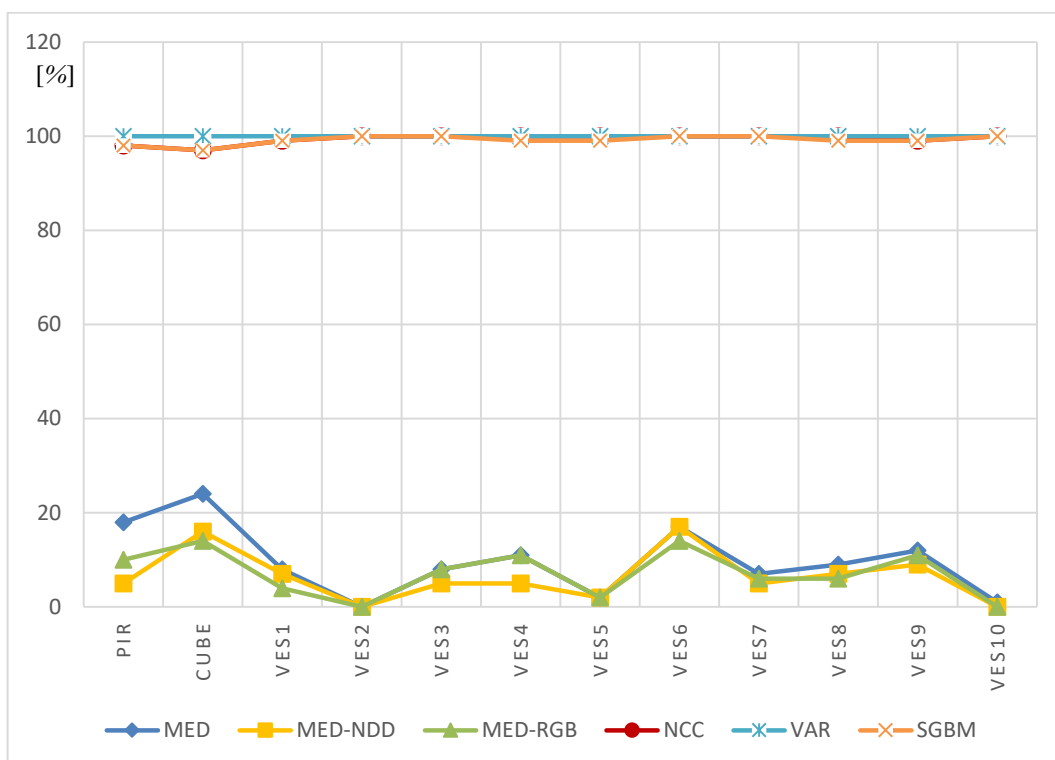
	Aloe	Baby1	Baby2	Baby3	Bowling1	Bowling2	Cloth1	Cloth2	Cloth3	Cloth4	Flowerpots	średnia
<b>MED</b>	85	79	70	86	91	91	80	100	94	80	88	86
<b>MED-NDD</b>	78	75	62	81	89	89	62	82	84	76	85	78
<b>MED-RGB</b>	77	74	71	81	87	89	70	82	84	76	83	79
NCC	45	54	52	60	74	73	25	72	48	62	74	58
VAR	<b>3</b>	7	25	29	62	38	8	20	10	20	41	24
SGBM	12	<b>6</b>	<b>6</b>	<b>12</b>	<b>19</b>	<b>16</b>	<b>2</b>	<b>17</b>	<b>3</b>	<b>4</b>	<b>21</b>	<b>11</b>

Tab. 9.46 Przedstawienie procentowej liczby błędnych sparowań  $\delta_D$  dla obrazów Middlebury – część 2.

	Lampshade1	Lampshade2	Mid1	Mid2	Monopoly	Plastic	Rocks1	Rocks2	Wood1	Wood2	średnia
<b>MED</b>	85	83	98	96	69	92	89	92	85	95	88
<b>MED-NDD</b>	84	80	81	74	60	87	74	66	83	83	77
<b>MED-RGB</b>	79	82	83	78	70	92	77	78	86	82	81
NCC	79	77	74	75	69	78	39	42	52	39	62
VAR	46	61	53	65	57	65	18	15	50	67	50
SGBM	<b>16</b>	<b>13</b>	<b>22</b>	<b>25</b>	<b>19</b>	<b>18</b>	<b>7</b>	<b>7</b>	<b>9</b>	<b>19</b>	<b>16</b>

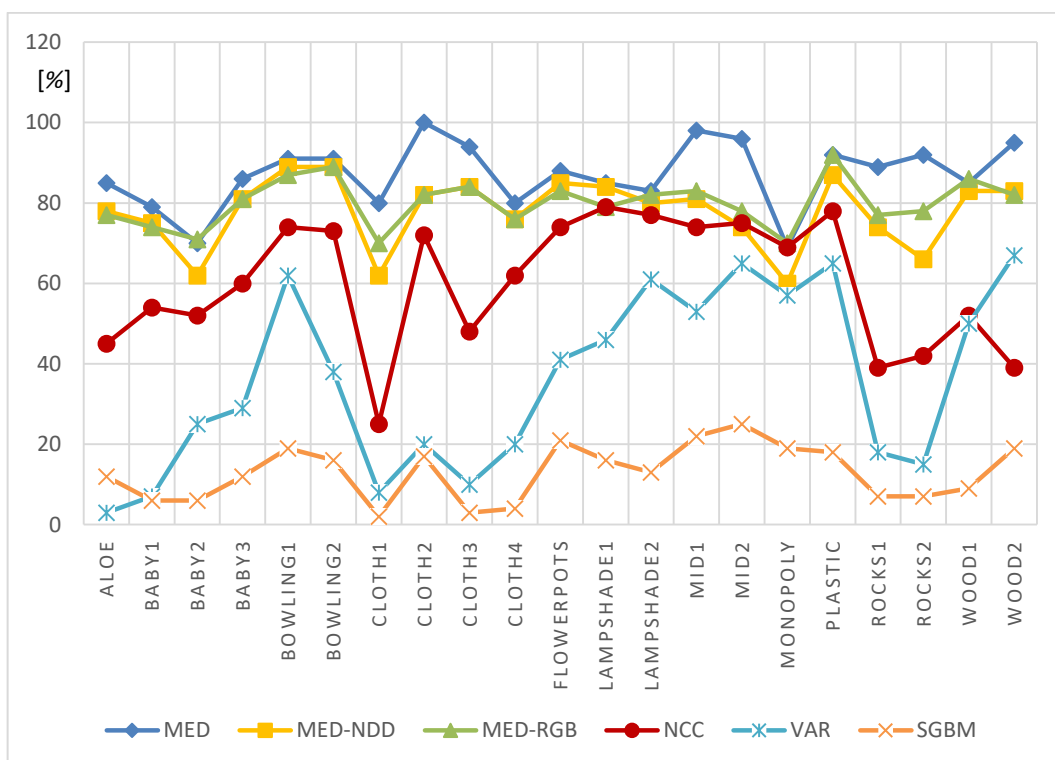
Porównując procentową liczbę błędnych sparowań  $\delta_D$  dla obrazów naczyń Ves1-Ves10 i modeli figur stereometrycznych najlepiej wypadają proponowane metody MED-NDD i MED-RGB osiągając najniższą wartość  $\delta_D$  bliską 0 % dla obrazu Ves2 i Ves10.

W ogólnym ujęciu algorytmy MED-NDD i MED-RGB osiągają wartości średnie  $\delta_D$  na takim samym poziomie 7 %.



Rys. 9.18. Procentowa liczba błędnych sparowań  $\delta_D$  dla obrazów własnych.

Pozostałe algorytmy parowania obrazów NCC, VAR, SGBM osiągają wartości  $\delta_D$  zawierające się od 98 % do 100 %.



Rys. 9.19. Procentowa liczba błędnych sparowań  $\delta_D$  dla obrazów Middlebury.

Dla obrazów testowych Middlebury najmniejszą wartość  $\delta_D = 2\%$  osiąga metoda SGBM dla obrazu Cloth1. Najniższą wartość średnią  $\delta_D = 17\%$  osiąga algorytm SGBM. Drugą najniższą średnią wartość  $\delta_D$  osiąga algorytm VAR z wartością  $37\%$ . Proponowane algorytmy MED, MED-NDD i MED-RGB osiągają tu wysokie wartości  $\delta_D$  na poziomie  $80\%$ .

W następnym podrozdziale przedstawione zostanie podsumowanie dyskusji.

## 9.4. Podsumowanie dyskusji

Z przedstawionych porównań wynika, iż **nowo opracowane** algorytmy MED, MED-NDD i MED-RGB podczas parowania obrazów wysegmentowanych naczyń:

- są średnio 7 tyś. razy szybsze, niż algorytm NCC, 125 razy szybsze, niż VAR i 4 razy szybsze, niż SGBM,
- osiągają 9 razy wyższą dokładność skuteczności parowania obrazów  $\delta_r$ , niż inne metody,
- osiągają 14 razy niższą wartość procentowej liczby błędnych sparowań  $\delta_D$ , niż inne metody.

Natomiast **znane** powszechnie stosowane algorytmy NCC, VAR i SGBM prezentują wysoką sprawność w parowaniu obrazów rzeczywistych z biblioteki Middlebury. Podczas parowania obrazów rzeczywistych z biblioteki Middlebury:

- są średnio 8 tyś. razy wolniejsze dla NCC, 200 razy wolniejsze dla VAR i 3.5 razy wolniejsze niż nowo proponowane algorytmy,
- osiągają średnio 2.5 razy wyższą dokładność skuteczności parowania obrazów  $\delta_r$ , niż nowo proponowane algorytmy,
- osiągają 7 razy niższą wartość procentowej liczby błędnych sparowań  $\delta_D$  w stosunku do nowo proponowanych algorytmów.

## 10. PODSUMOWANIE

Celem pracy było stworzenie algorytmu pozwalającego na znalezienie odpowiedniości punktów na obrazach stereowizyjnych. Założono również wyprofilowanie algorytmu do zastosowań medycznych, a jako dziedzinę algorytmu wybrano obrazy binarne

Opracowano trzy metody segmentacji obrazów (BGOM, SSSB, BLG).

- Pierwsza metoda (BGOM) ekstrahuje osie naczyń podskórnych. Eliminacja powierzchni homogenicznych i zawężenie treści obrazu ułatwia poszukiwanie na nich odpowiedniości. Metoda jest dedykowana do segmentacji obrazów medycznych.
- Druga metoda (SSSB) została stworzona na potrzeby badania wpływu cech na jakość stereokorespondencji, a wynik jej działania jest podobny do obrazów otrzymanych w wyniku działania operatora Canny'ego. Metodę cechuje szybkość działania oraz możliwość kontroli nad cechami obrazu. Podobnie jak poprzednia metoda, jest ona dedykowana do segmentacji zobrażeń medycznych.
- Trzecia metoda segmentacji (BLG), lepiej się sprawdza do segmentacji rzeczywistych zobrażeń. Została ona opracowana, by przetestować algorytmy poszukiwania odpowiedniości na obrazach rzeczywistych udostępnionych przez Middlebury College, które pozostają bazą odniesienia dla testowania wszystkich nowych metod znajdowania odpowiedniości.

Opracowano trzy metody poszukiwania odpowiedniości dla obrazów binarnych (MED, MED-NDD, MED-RGB).

- Pierwsza metoda (MED) bazuje na założeniu, że prawidłową dysparycją jest minimalna odległość pomiędzy punktem na obrazie lewym i wszystkimi punktami na obrazie prawym dla zobrażeń zrektyfikowanych.
- Druga metoda (MED-NDD), uzgadnia poprawność dysparycji obliczonych zgodnie z metodą pierwszą, dodatkowo sprawdzając warunek odległości do sąsiednich punktów, które powinny być zbliżone pomiędzy obrazem lewym i prawym. w ten sposób poprawiono wyniki o blisko 30 % dla zobrażeń medycznych i blisko 8 % dla obrazów z Middlebury, w stosunku do pierwszej metody, jednocześnie wydłużając czas pracy algorytmu dwukrotnie dla zobrażeń medycznych i pięciokrotnie dla obrazów Middlebury.

- Trzecia metoda (MED-RGB), uzgadnia poprawność dysparycji obliczonych zgodnie z metodą pierwszą, sprawdzając dodatkowo warunek podobieństwa składowych barwy dla pikseli otaczających punkty homologiczne na stereoparze. Efektywność tej metody jest zbliżona do metody drugiej.

W zależności od zobrażeń czasami druga metoda prezentuje lepszą efektywność, a czasami trzecia (Rys. 7.4). Trzecia metoda jest trzykrotnie wolniejsza od pierwszej dla obrazów medycznych i aż trzynastokrotnie wolniejsza dla obrazów Middlebury.

Przeprowadzono ocenę dokładności powstałej w wyniku poszukiwania odpowiedniości macierzy dysparycji przyjmując trzy kryteria dokładnościowe oraz bazując na wzorcowych macierzach dysparycji.

Dokonano oceny wpływu doboru cech obrazu na jakość wyników poszukiwania odpowiedniości punktów charakterystycznych na obrazach stereowizyjnych. z obrazów pozyskano cechy znaczące i wykazano, iż wybrane cechy, mają znaczący wpływ na jakość wyników odpowiedniości punktów charakterystycznych na obrazach stereowizyjnych. Wykazano również, iż sterując cechami znaczącymi można wpłynąć na polepszenie wyników poszukiwania odpowiedniości (Tab. 8.34). Na podstawie otrzymanych wyników można wyprowadzić wniosek o prawidłowym doborze cech znaczących. Jednak by właściwie sterować cechami należało by przeprowadzić dalsze badania na odpowiednio do tego przygotowanych obrazach.

Dzięki wyprofilowaniu algorytmów poszukiwania odpowiedniości do pracy na obrazach naczyń podskórnych, prezentują one znacznie wyższą sprawność w parowaniu obrazów binarnych naczyń podskórnych w stosunku do wyników, jakie osiągają znane algorytmy NCC, VAR i SGBM, co świadczy o odpowiednim zoptymalizowaniu algorytmów do parowania obrazów medycznych.

Zaproponowane algorytmy MED, MED-NDD i MED-RGB należą do najszybciej działających wśród testowanych algorytmów.

Opierając się na powyższym, należy stwierdzić, iż przeprowadzone badania potwierdzają słuszność sformułowanych na początku pracy tez, iż dedykowane metody analizy i przetwarzania obrazów stereowizyjnych pozwalają na polepszenie jakości otrzymywanych wyników stereokorespondencji oraz, iż wprowadzenie nowych cech punktów charakterystycznych pozwala na podniesienie skuteczności działania algorytmów poszukiwania odpowiedniości.

W dalszej swojej pracy autor niniejszej rozprawy chciałby przede wszystkim kontynuować badania nad segmentacją zobrażeń medycznych oraz algorytmami



poszukiwania odpowiedniości. z przeprowadzonych badań wynika, iż jakość segmentacji znacząco wpływa na wyniki poszukiwania odpowiedniości. Przeprowadzone badania autora pozwalają na stwierdzenie, że dla każdego rodzaju zobrazowania medycznego konieczne jest opracowanie dedykowanej metody segmentacji, a później również metody poszukiwania odpowiedniości punktów charakterystycznych, odpowiednio sprofilowanej do charakteru reprezentacji binarnej.

Opracowane metody segmentacji i parowania obrazów pozwalają na zastosowanie ich w stacji Blootbot opracowanej przez Imperial College London, która jak do tej pory korzysta z techniki detekcji naczyń opartej na analizie siły nacisku igły. Zaimplementowanie głowicy stereowizyjnej oraz iluminatora naczyniowego razem z opracowanymi algorytmami segmentacji i parowania obrazów pozwoliło by na daleko posuniętą modernizację stacji Blootbot. Dlatego autor będzie starał się nawiązać współpracę z twórcami Blootbot'a celem jego dalszego rozwoju.

Planuje się prowadzić dalsze badania celem jednoznacznego ustalenia uniwersalnych wartości cech znaczących pozwalających osiągać minimalne wartości błędów skuteczności parowania oraz liczby błędnych sparowań. w tym celu, w pierwszej kolejności planuje się stworzyć dedykowaną bazę obrazów własnych. Bardzo ciekawą implementacją opracowanych metod segmentacji i parowania obrazów mogło by być monitorowanie procesów rehabilitacji po przebytych zabiegach np. alloplastyki stawu biodrowego, czy stawu kolanowego. Autor niniejszej rozprawy bazując na już nawiązanej współpracy z Oddziałem Chirurgii Urazowo - Ortopedycznej Szpitala w Chorzowie będzie chciał zająć się opracowaniem metody monitorowania rehabilitacji, a następnie jej wdrożeniem.

## BIBLIOGRAFIA

- [1] Kurzynski, M., Krysmann, M., Trajdos, P., Wolczowski, A.: *Multiclassifier system with hybrid learning applied to the control of bioprosthetic hand*. Computers in Biology and Medicine. 69, 286–297 (2016).
- [2] Kurzynski, M., Majak, M.: *Meta-Bayes Classifier with Markov Model Applied to the Control of Bioprosthetic Hand*. In: Smart Innovation, Systems and Technologies. pp. 107–117. Springer Verlag (2016).
- [3] Tokarczyk, R.: Fotogrametryczne pomiary geometrii ciała ludzkiego w zastosowaniu do badania wad postawy. AGH Uczelniane Wydawnictwa Naukowo-Dydaktyczne, Kraków (2009).
- [4] Härtl, R., Lam, K.S., Wang, J., Korge, A., Kandziora, F., Audigé, L.: *Worldwide survey on the use of navigation in spine surgery*. World Neurosurgery. 79, 162–172 (2013).
- [5] Benjamin J. Shin, Andrew R. James, Innocent U. Njoku, Roger Härtl: *Pedicle screw navigation: a systematic review and meta-analysis of perforation risk for computer-navigated versus freehand insertion*. Journal of Neurosurgery: Spine. 17, 113–122 (2012).
- [6] D'Apuzzo, N.: *Measurement and modeling of human faces from multi images*. International Archives of Photogrammetry and Remote Sensing. 34 (5), 241–246 (2002).
- [7] Tokarczyk, R., Wróbel, Z., Huppert, M.: Automatyzacja pomiaru na obrazach cyfrowych w systemie fotogrametrycznym do badania wad postawy. AGH Uczelniane Wydawnictwa Naukowo-Dydaktyczne, Kraków (2007).
- [8] Malian, A., Azizi, A., Van Den Heuvel, F. a.: *Medphos: a new photogrammetric system for medical measurement*. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 35, 929–933 (2004).
- [9] Mert, A., Gan, L.S., Knosp, E., Sutherland, G.R., Wolfsberger, S.: *Advanced cranial navigation*. Neurosurgery. 72, (2013).
- [10] Bandiera, S., Ghermandi, R., Gasbarrini, A., Colangeli, S., Boriani, S.: *Navigation-assisted surgery for tumors of the spine*. European Spine Journal. 22, (2013).
- [11] Maciejczak, A.: *Frameless image-guided brain and spine surgery | Nawigacja w chirurgii mózgu i kregoslupa*. Przegląd lekarski. 72, (2015).
- [12] Kim, J.-Y., Cheon, S.-H., Kyung, H.-S.: *Mid-term Results of Total Knee Arthroplasty Using PFC Sigma RP-F*. Knee surgery & related research. 24, 221–6 (2012).
- [13] Synder, M., Altimimi, M.A.H., Borowski, A., Sibiński, M., Drobniowski, M.: *Evaluation of Outcomes of Total Knee Replacement with and without a Navigation System*. Ortopedia Traumatologia Rehabilitacja. 18, 251–261 (2016).
- [14] Renkawitz, T., Tingart, M., Grifka, J., Sendtner, E., Kalteis, T.: *Computer-assisted total hip arthroplasty: coding the next generation of navigation systems for orthopedic surgery*. Expert Rev Med Devices. 6, 507–514 (2009).
- [15] Burduk, P., Dalke, K., Kaźmierczak, W.: *Nawigacja śródoperacyjna w chirurgii endoskopowej zatok przynosowych*. Otolaryngologia Polska. 66, 36–39 (2012).
- [16] Citardi, M.J., Batra, P.S.: *Intraoperative surgical navigation for endoscopic sinus surgery: rationale and indications*. Current opinion in otolaryngology & head and neck surgery. 15, 23–27 (2007).
- [17] Cuper, N.J., de Graaff, J.C., Verdaasdonk, R.M., Kalkman, C.J.: *Near-Infrared Imaging in Intravenous Cannulation in Children: A Cluster Randomized Clinical Trial*. Pediatrics. 131, e191–e197 (2013).
- [18] Graham-Rowe, D.: *A robot in the right vein*. New Scientist. (1999).
- [19] Sitek, Z.: Fotogrametria ogólna i inżynierska. PPWK, Warszawa (1991).
- [20] Wang, J., Shi, F., Zhang, J., Liu, Y.: *A new calibration model of camera lens distortion*.

- Pattern Recognition. 41, 607–615 (2008).
- [21] Brown, D.C.: *Decentering Distortion of Lenses*. Photometric Engineering. 32, 444–462 (1966).
- [22] Heikkila, J., Silven, O.: *A four-step camera calibration procedure with implicit image correction*. In: Proceedings of IEEE Computer Society, Conference on Computer Vision and Pattern Recognition. pp. 1106–1112. IEEE Comput. Soc (1997).
- [23] Zhang, Z.Z.Z.: *Flexible camera calibration by viewing a plane from unknown orientations*. Proceedings of the Seventh IEEE International Conference on Computer Vision. 1, 0–7 (1999).
- [24] Zhang, Z.: *A flexible new technique for camera calibration*. IEEE Transactions on Pattern Analysis and Machine Intelligence. 22, 1330–1334 (2000).
- [25] Cyganek, B.: Komputerowe przetwarzanie obrazów trójwymiarowych. Exit, Warszawa (2002).
- [26] Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press (2003).
- [27] Luong, Q.-T., Faugeras, O.D.: *The fundamental matrix: Theory, algorithms, and stability analysis*. International Journal of Computer Vision. 17, 43–75 (1996).
- [28] Bouguet, J.-Y.: *Complete Camera Calibration Toolbox for Matlab*, <http://www.vision.caltech.edu/bouguetj/>. Dostępne na 15.01.2017.
- [29] Brown, M.Z., Burschka, D., Hager, G.D.: *Advances in computational stereo*, (2003).
- [30] Cyganek, B., Siebert, J.P.: An Introduction to 3D Computer Vision Techniques and Algorithms. (2009).
- [31] **Popielski, P.**, Wróbel, Z.: *The Feature Detection on the Homogeneous Surfaces with Projected Pattern*. In: Piętka, E. and Kawa, J. (eds.) Information Technologies in Biomedicine. pp. 118–128. Lecture Notes in Computer Science, Springer Berlin Heidelberg (2012).
- [32] Harris, C., Stephens, Mike: *A Combined Corner and Edge Detector*. Proceedings of the Alvey Vision Conference 1988. 147–151 (1988).
- [33] Noble, J.A.: *Descriptions of image surfaces*. PhD thesis, Department of Engineering Science. (1989).
- [34] Lewis, J.P.: *Fast normalized cross-correlation*. Vision Interface. 120–123 (1995).
- [35] Konolige, K.: *Small Vision Systems: Hardware and Implementation*. Robotics Research. 203–212 (1998).
- [36] Birchfield, S., Tomasi, C.: *Depth discontinuities by pixel-to-pixel stereo*. International Journal of Computer Vision. 35, 269–293 (1999).
- [37] Kosov, S., Thormählen, T., Seidel, H.P.: *Accurate real-time disparity estimation with variational methods*. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 5875 LNCS, 796–807 (2009).
- [38] Kolmogorov, V., Zabih, R.: *Computing Visual Correspondence with Occlusions via Graph Cuts*. In: Proc. IEEE International Conference on Computer Vision (ICCV). pp. 508–515 (2001).
- [39] Zhang, Z., Deriche, R., Faugeras, O., Luong, Q.T.: *A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry*. Artificial Intelligence. 78, 87–119 (1995).
- [40] Gong, M.: *Multi-resolution stereo matching using genetic algorithm*. In: Proceedings IEEE Workshop on Stereo and Multi-Baseline Vision (SMBV 2001). pp. 21–29 (2001).
- [41] Cyganek, B., Korohoda, P., Borgosz, J.: *Matching Stereo-Images with Cellular Neural Networks*. In: Fourth Conference ‘Neural Networks and Their Applications’., Zakopane (1999).

- 
- [42] Cyganek, B.: *Neural Networks Application to The Correlation-Based Stereo-Images Matching*. In: Engineering Applications of Neural Networks, Proceedings of the 5th International Conference EANN '99, Warsaw, Poland. pp. 17–22 (1999).
- [43] Scharstein, D., Szeliski, R.: *Stereo Matching with Nonlinear Diffusion*. International Journal of Computer Vision. 28, 155–174 (1998).
- [44] Sun, J., Zheng, N.N., Shum, H.Y.: *Stereo matching using belief propagation*. IEEE Transactions on Pattern Analysis and Machine Intelligence. 25, 787–800 (2003).
- [45] Hirschmüller, H.: *Stereo processing by semiglobal matching and mutual information*. IEEE Transactions on Pattern Analysis and Machine Intelligence. 30, 328–341 (2008).
- [46] Pastuszek, W.: Barwa w grafice komputerowej. PWN, Warszawa (2000).
- [47] Woźniak, W.: *Kolorymetria i widzenie barwne*, <http://www.if.pwr.edu.pl/~wozniak/kolorymetria.html>. Dostępne na 15.01.2017.
- [48] Matthews, Gary, G.: Neurobiologia. Od cząsteczek i komórek do układów. PZWL, Warszawa (2000).
- [49] Milner, David, A., Goodale, Melvyn, A.: Mózg wzrokowy w działaniu. Wydawnictwo Naukowe PWN, Warszawa (2008).
- [50] Jianbo Shi, C. Tomasi: *Good features to track*. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition CVPR-94. pp. 593–600 (1994).
- [51] Tomasi, C., Kanade, T.: Detection and Tracking of Point Features. (1991).
- [52] Oskoei, M.A., Hu, H.: *A Survey on Edge Detection Methods*. Evaluation. 1–36 (2010).
- [53] Shrivakshan, G.T., Chandrasekar, C.: *A Comparison of various Edge Detection Techniques used in Image Processing*. International Journal of Computer Science Issues (IJCSI). 9, 269–276 (2012).
- [54] McIlhagga, W.: *The canny edge detector revisited*. International Journal of Computer Vision. 91, 251–261 (2011).
- [55] Mukhopadhyay, P., Chaudhuri, B.B.: *A survey of Hough Transform*. Pattern Recognition. 48, 993–1010 (2015).
- [56] Żorski, W.: Metody segmentacji obrazów oparte na transformacie Hough’a. Instytut Automatyki i Robotyki, Wydział Cybernetyki WAT, Warszawa (2000).
- [57] Cyganek, B., Siebert, J.P.: An Introduction to 3D Computer Vision Techniques and Algorithms. John Wiley & Sons (2011).
- [58] Olofsson, A.: *Modern stereo correspondence algorithms: investigation and evaluation*. PhD diss., Master Thesis, Department of electrical engineering, Linköpings Universitet. 114 (2010).
- [59] Meijering, E.: *A chronology of interpolation: From ancient astronomy to modern signal and image processing*. Proceedings of the IEEE. 90, 319–342 (2002).
- [60] Brox, T., Bruhn, A., Papenberg, N., Weickert, J.: *High accuracy optical flow estimation based on a theory for warping*. Computer Vision-ECCV 2004, Lecture Notes in Computer Science. 3024, 25–36 (2004).
- [61] Kosov, S.: *3D map reconstruction with variational methods*. Master thesis, Saarland University. (2008).
- [62] Papenberg, N., Bruhn, A., Brox, T., Didas, S., Weickert, J.: *Highly accurate optic flow computation with theoretically justified warping*. International Journal of Computer Vision. 67, 141–158 (2006).
- [63] Hirschmüller, H.: *Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information*. In: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on (Volume: 2). pp. 807–814 (2005).
- [64] Boykov, Y., Veksler, O., Zabih, R.: *Efficient Approximate Energy Minimization via Graph Cuts*. IEEE transactions on Pattern Analysis and Machine Intelligence. 20, 1222–1239 (2001).

- 
- [65] Van Meerbergen, G., Vergauwen, M., Pollefeys, M., Van Gool, L.: *A hierarchical symmetric stereo algorithm using dynamic programming*. International Journal of Computer Vision. 47, 275–285 (2002).
- [66] Chang, Y.-C., Datchev, I., Habib, A.: *A Photogrammetric System for 3D Reconstruction of a Scoliotic Torso*. Proceedings of the ASPRS 2009 Annual Conference. (2008).
- [67] Bailey, D.H., Borwein, J.M., Mattingly, A., Wightwick, G.: *The computation of previously inaccessible digits of  $\pi^2$  and Catalan's constant*. (2013).
- [68] **Popielski, P.**, Wróbel, Z.: *An attempt to optimize the process of automatic point matching for homogeneous surface objects*. Archiwum Fotogrametrii, Kartografii i Teledetekcji. 22, 351–361 (2011).
- [69] **Popielski, P.**, Wróbel, Z.: *Problemy rekonstrukcji 3D ze zdjęć niometrycznych w paśmie widzialnym w zastosowaniach medycznych*. In: Systemy wspomagania decyzji. pp. 259–266. Uniwersytet Śląski, Katowice (2012).
- [70] **Popielski, P.**, Wróbel, Z., Koprowski, R.: *Object Detail Correspondence Problem in Stereovision*. In: Piętka, E., Kawa, J., and Wieclawek, W. (eds.) Information Technologies in Biomedicine, Volume 3. pp. 209–222. Advances in Intelligent Systems and Computing, Springer International Publishing (2014).
- [71] Derpanis, K.G.: *Overview of the RANSAC Algorithm*. Image Rochester NY. 4, 2–3 (2010).
- [72] Kraus, K.: *Photogrammetry: Geometry from Images and Laser Scans*. Walter de Gruyter (2007).
- [73] **Popielski, P.**, Koprowski, R., Wróbel, Z.: *The fast matching algorithm for rectified stereo images*. In: Advances in Intelligent Systems and Computing. pp. 107–118. Advances in Intelligent Systems and Computing, Springer International Publishing (2016).
- [74] **Popielski, P.**, Koprowski, R., Wróbel, Z., Wilczyński, S., Doroz, R., Wróbel, K., Porwik, P.: *The Matching Method for Rectified Stereo Images Based on Minimal Element Distance and RGB Component Analysis*. In: Nguyen, N.T., Iliadis, L., Manolopoulos, Y., and Trawiński, B. (eds.) Computational Collective Intelligence: 8th International Conference, ICCCI 2016, Halkidiki, Greece, September 28–30, 2016. Proceedings, Part II. pp. 482–493. Springer International Publishing, Cham (2016).
- [75] Scharstein, D., Szeliski, R.: *High-accuracy stereo depth maps using structured light*. 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings. 1, 195–202 (2003).
- [76] Scharstein, D., Pal, C.: *Learning conditional random fields for stereo*. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. (2007).
- [77] Hirschmuller, H., Scharstein, D.: *Evaluation of Cost Functions for Stereo Matching*. Proc. of CVPR. 1–8 (2007).
- [78] Zhang, J., Hu, J.: *Image Segmentation Based on 2D Otsu Method with Histogram Analysis*. 2008 International Conference on Computer Science and Software Engineering. 6, 105–108 (2008).
- [79] Sauvola, J., Pietikäinen, M.: *Adaptive document image binarization*. Pattern Recognition. 33, 225–236 (2000).
- [80] Scharstein, D., Szeliski, R.: *A taxonomy and evaluation of dense two-frame stereo correspondence algorithms*. International Journal of Computer Vision. 47, 7–42 (2002).
- [81] Yang, Q., Wang, L., Yang, R., Stewénius, H., Nistér, D.: *Stereo matching with color-weighted correlation, hierarchical belief propagation, and occlusion handling*. IEEE Transactions on Pattern Analysis and Machine Intelligence. 31, 492–504 (2009).
- [82] Wang, Z.-F., Zheng, Z.-G.: *A region based stereo matching algorithm using cooperative*

- optimization*. Proc. Conf. Computer Vision Pattern Rec. (CVPR). 0, 1–8 (2008).
- [83] Papadakis, N., Caselles, V.: *Multi-label depth estimation for graph cuts stereo problems*. Journal of Mathematical Imaging and Vision. 38, 70–82 (2010).
  - [84] Yang, Q., Wang, L., Yang, R., Wang, S., Liao, M., Nister, D.: *Real-time Global Stereo Matching Using Hierarchical Belief Propagation*. Bmvc. 6, 989–998 (2006).
  - [85] Wang, L., Liao, M., Gong, M., Yang, R., Nister, D.: *High-quality real-time stereo using adaptive cost aggregation and dynamic programming*. In: Proceedings - Third International Symposium on 3D Data Processing, Visualization, and Transmission, 3DPVT 2006. pp. 798–805 (2007).
  - [86] Olague, G., Fernández, F., Pérez, C.B., Lutton, E.: *The infection algorithm: an artificial epidemic approach for dense stereo correspondence*. Artificial life. 12, 593–615 (2006).
  - [87] Klaus, A., Sormann, M., Karner, K.: *Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure*. In: Proceedings - International Conference on Pattern Recognition. pp. 15–18 (2006).
  - [88] Bleyer, M., Gelautz, M.: *A layered stereo matching algorithm using image segmentation and global visibility constraints*. ISPRS Journal of Photogrammetry and Remote Sensing. 59, 128–150 (2005).
  - [89] Strecha, C., Fransens, R., Van Gool, L.: *Combined depth and outlier estimation in multi-view stereo*. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. pp. 2394–2401 (2006).
  - [90] Hirschmuller, H., Scharstein, D.: *Evaluation of stereo matching costs on images with radiometric differences*. IEEE Transactions on Pattern Analysis and Machine Intelligence. 31, 1582–1599 (2009).
  - [91] **Popielski, P.**, Wrobel, Z., Koprowski, R.: *The Effectiveness of Matching Methods for Rectified Images*. In: Burduk, R., Jackowski, K., Kurzynski, M., Wozniak, M., and Zolnierrek, A. (eds.) Proceedings of the 8th International Conference on Computer Recognition Systems CORES 2013. pp. 479–489. Advances in Intelligent Systems and Computing, Springer International Publishing (2013).